



University of Pennsylvania
ScholarlyCommons

Technical Reports (CIS)

Department of Computer & Information Science

July 2002

Integral Transforms and Signal Processing on the 2-Sphere

Daniel Rudoy
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/cis_reports

Recommended Citation

Daniel Rudoy, "Integral Transforms and Signal Processing on the 2-Sphere", . July 2002.

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-02-20.

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/cis_reports/175
For more information, please contact repository@pobox.upenn.edu.

Integral Transforms and Signal Processing on the 2-Sphere

Abstract

This paper is concerned with the group-theoretical background of integral transforms and the role they play in signal processing on the sphere. An overview of topological groups, measure and representation theories, and an overview of the Windowed Fourier Transform and the Continuous Wavelet Transform are presented. A group-theoretical framework within which these transforms arise is presented. The connection between integral transforms and square-integrable group representations is explored. The theory is also generalized beyond groups to homogeneous spaces. The abstract theory is then applied to signal processing on the sphere with a discussion of both global and local methods. A detailed derivation of the continuous spherical harmonic transform is presented. Global methods such as the spherical Fourier transform and convolution are presented in an appendix as well as some background material from group theory and topology.

Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-02-20.

Integral Transforms and Signal Processing on the 2-Sphere

Daniel Rudoy^{*†}

Department of Computer and Information Science
University of Pennsylvania, Philadelphia, USA
rudoy@grasp.cis.upenn.edu

July 8, 2002

Technical Report: MS-CIS-02-20

Abstract

This paper is concerned with the group-theoretical background of integral transforms and the role they play in signal processing on the sphere. An overview of topological groups, measure and representation theories, and an overview of the Windowed Fourier Transform and the Continuous Wavelet Transform are presented. A group-theoretical framework within which these transforms arise is presented. The connection between integral transforms and square-integrable group representations is explored. The theory is also generalized beyond groups to homogeneous spaces. The abstract theory is then applied to signal processing on the sphere with a discussion of both global and local methods. A detailed derivation of the continuous spherical harmonic transform is presented. Global methods such as the spherical Fourier transform and convolution are presented in an appendix as well as some background material from group theory and topology.

^{*}The author is grateful for support through the following grants: NSF-IIS-0083209, NSF-EIA-0218691, NSF-IIS-0121293, NSF-EIA-9703220, a DARPA/ITO/NGI subcontract to UNC.

[†]This work was completed as a Masters Thesis at the Computer and Information Science Department, University of Pennsylvania in July, 2002.

Contents

Introduction	2
Mathematical Background	3
1 Hilbert Spaces	3
2 Locally Compact Groups	7
3 $SO(3)$ and the 2-Sphere	8
3.1 Stereographic Projection	10
Basic Measure Theory	12
1 On the use of Measure Theory	12
2 Measure Spaces	13
3 Abstract Integration	15
4 Haar Measure	16
Basic Representation Theory	19
1 On the use of representations	19
2 Definition of a Representation	19
3 Representations of Finite Groups	20
4 Representations of Locally Compact Groups	23
Integral Transforms	26
1 General Construction	27
2 Extension to Homogeneous Spaces	32

Examples of Integral Transforms	35
1 Continuous Wavelet Transform on \mathbb{R}	35
2 Windowed Fourier Transform	38
3 Continuous Wavelet Transform on \mathbb{R}^2 and beyond	40
Signal Processing on the 2-Sphere	42
1 Introduction	42
2 Wavelet Transform on the 2-Sphere	43
2.1 Preliminaries	43
2.2 The Construction	45
Conclusions and Future Work	48
Appendix A: Elementary Group Theory	49
Appendix B: Elementary Topology	52
Appendix C: Global Spherical Signal Processing	54
1 Spherical Fourier Transform	54
2 Convolution	56
Bibliography	59

Introduction

Over the last few decades the field of signal processing has weathered a revolution. Techniques that previously dominated the field such as Fourier transforms now have to compete with many other integral transforms and in particular, wavelet transforms. Moreover, there exists a number of different generalizations of each type of transform to higher dimensions and to different domains. A newcomer to signal processing might not only be overwhelmed by the abundance of various techniques, but may also be concerned with whether or not there are other techniques to be uncovered. From a theoretical standpoint this is quite a natural question to ask. The goal in this paper is to try and address both issues.

The main focus of this paper will be to show that there is a unifying, group-theoretical framework that underlies many of the integral transforms that one finds in signal processing today. We will describe the general theory and explicitly show how some integral transforms arise in its context. Specifically, we will demonstrate how to derive Continuous Wavelet Transforms and Windowed Fourier Transforms that operate on certain spaces of square-integrable functions. Hopefully, seeing these transforms developed within the same framework will allow the reader to study them and others in a well-organized and coherent manner.

The presentation of all of the material in this paper is reasonably self-contained and assumes a background in basic undergraduate mathematics. All of the other tools used, such as measure and representation theories, are developed from scratch. This enables us to present a general theory of integral transforms without hiding many of the details. This is done in the hope that the reader, once familiarized with the construction presented, will be able to design and apply novel integral transforms.

While not all possible integral transforms can be derived from the framework that is presented, a large number of cases is covered. Nevertheless, should the reader be interested in studying transforms that are not covered by the general theory presented here, he will find the material presented here quite useful for their study.

Mathematical Background

In this chapter, we review a number of concepts which we will use from time to time, ranging from the theory of Hilbert Spaces to some facts about locally compact groups. A number of relevant results are included without proof. It is assumed that the reader is familiar with basic algebra, analysis and topology. Nonetheless, some definitions are reviewed in Appendices A and B.

1 Hilbert Spaces

Definition 1.1. A *normed linear space* is a linear space in which to each vector \mathbf{x} there is a corresponding real number, denoted by $\|\mathbf{x}\|$ and called the norm, such that:

1. $\|\mathbf{x}\| \geq 0$, and $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = 0$
2. $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

In the above definition, α belongs to the field over which the linear space is defined. In general, we leave the field of scalars unspecified. Normed linear spaces can be viewed as metric spaces with respect to the metric defined by $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$. Considering normed linear spaces from this point of view turns out to be quite useful and motivates the following definition.

Definition 1.2. A *Banach Space* is a complete normed linear space with respect to the metric defined by $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$.

Example 1.1. Classic examples of Banach Spaces are $\mathcal{L}^p(\mathbf{X})$ spaces, composed of real-valued functions on a set \mathbf{X} with finite norm, defined by

$$\|\mathbf{f}\|_p = \left(\int_{\mathbf{X}} |f(x)|^p dx \right)^{\frac{1}{p}}$$

However, Banach Spaces are still somewhat too general to be useful for our purposes. Much of our work will focus on function spaces with additional structure. In particular, we require function spaces that are endowed with some geometrical property that enables us to discuss relative orientations of vectors in these spaces. One such property, orthogonality, will be crucial in the context of discussing integral transforms. Thus motivated, we come to the notion of a Hilbert Space as a specialized Banach Space.

Definition 1.3. A *Hilbert Space* \mathcal{H} is a Banach Space whose norm arises from a complex valued function of two vectors denoted by (\mathbf{x}, \mathbf{y}) . This function is called the inner product and, given $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{H}$, satisfies the three properties listed below.

1. $(\mathbf{x}, \mathbf{x}) = \|\mathbf{x}\|^2$
2. $(\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}) = \alpha(\mathbf{x}, \mathbf{z}) + \beta(\mathbf{y}, \mathbf{z})$
3. $\overline{(\mathbf{x}, \mathbf{y})} = (\mathbf{y}, \mathbf{x})$

Example 1.2. \mathbb{R}^n and \mathbb{C}^n are finite-dimensional Hilbert spaces, with the standard dot product of vectors as the inner product.

Example 1.3. The space of real-valued square-integrable functions on the real line, $\mathcal{L}^2(\mathbb{R})$, is an infinite-dimensional Hilbert Space. The inner product (\mathbf{f}, \mathbf{g}) is defined by:

$$(\mathbf{f}, \mathbf{g}) = \int_{-\infty}^{\infty} f(x)g(x)dx$$

The inner product imposes additional geometric structure on Banach Spaces. In particular, an inner product can be used to define the notion of orthogonality for two vectors. From now on, let \mathcal{H} denote a Hilbert Space unless otherwise specified.

Definition 1.4. If $(\mathbf{x}, \mathbf{y}) = 0$ for two vectors $\mathbf{x}, \mathbf{y} \in \mathcal{H}$, then \mathbf{x}, \mathbf{y} are called *orthogonal vectors*.

Definition 1.5. Let U be a subspace of \mathcal{H} . We say that a vector $\mathbf{x} \in \mathcal{H}$ is *perpendicular* to U if for all $\mathbf{y} \in U$, $(\mathbf{x}, \mathbf{y}) = 0$, and we write $\mathbf{x} \perp U$.

Definition 1.6. Let U, V be subspaces of a Hilbert Space \mathcal{H} . We say that U and V are *orthogonal subspaces* if for all $\mathbf{x} \in U$, $\mathbf{x} \perp V$, and we write $U \perp V$.

To conclude our discussion of orthogonality, we introduce direct sums. Later, they will be used to define irreducible representations. For a more in depth discussion of direct sums, see Artin [18].

Definition 1.7. Let V_1, V_2, \dots, V_n be pairwise orthogonal subspaces of \mathcal{H} . Let $W = \{w \in W \mid w = v_1 + \dots + v_n, v_i \in V_i\}$. We say that W is the *direct sum* of V_1, \dots, V_n , and we write: $W = V_1 \oplus \dots \oplus V_n$.

The goal of the next few definitions will be to formalize the notion of operators on Hilbert Spaces. They will be crucial in the study of wavelet transforms and frames. We begin with the definition of a linear transformation.

Definition 1.8. Let \mathcal{H} and \mathcal{H}' be Hilbert spaces and let α be a scalar. A mapping T of \mathcal{H} into \mathcal{H}' is called a *linear transformation* if:

1. $T(\alpha \mathbf{x}) = \alpha T(\mathbf{x})$
2. $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$

If we view a linear transformation of two linear spaces as a mapping between metric spaces, it is natural to wonder what properties this mapping has. Two usual properties to investigate are continuity and boundedness.

Definition 1.9. Let \mathcal{H} and \mathcal{H}' be Hilbert spaces and let T be a linear transformation of \mathcal{H} into \mathcal{H}' . We say that T is *continuous*, when it is continuous as a mapping of the metric space \mathcal{H} into the metric space \mathcal{H}' . In other words as $\mathbf{x}_n \rightarrow \mathbf{x}$ in \mathcal{H} , $T(\mathbf{x}_n) \rightarrow T(\mathbf{x})$ in \mathcal{H}' .

Definition 1.10. Let \mathcal{H} and \mathcal{H}' be Hilbert spaces and let T be a linear transformation of \mathcal{H} into \mathcal{H}' . We say that T is *bounded* if for all $\mathbf{x} \in \mathcal{H}$ there exists a real number $k \geq 0$ such that:

$$\|T(\mathbf{x})\| \leq k\|\mathbf{x}\|$$

The most important idea we have to keep in mind is that continuous and bounded linear transformations on Hilbert Spaces are one and the same.

Theorem 1.1. Let \mathcal{H} and \mathcal{H}' be Hilbert spaces and let T be a linear transformation of \mathcal{H} into \mathcal{H}' . Then T is continuous iff T is bounded.

Proof. Assume to the contrary that T is continuous and that there does not exist $k \in \mathbb{R}$ such that $\|T(\mathbf{x})\| \leq k\|\mathbf{x}\|$. Then for each positive integer n we can find a vector $\mathbf{x}_n \in \mathcal{H}$ such that $\|T(\mathbf{x}_n)\| > n\|\mathbf{x}_n\|$. Equivalently, $\|T(\mathbf{x}_n/n\|\mathbf{x}_n\|)\| > 1$. (The strict inequality prohibits the case when $\|\mathbf{x}_n\| = 0$). Let $\mathbf{y}_n = \mathbf{x}_n/n\|\mathbf{x}_n\|$. It is easy to see that while $\mathbf{y}_n \rightarrow 0$, $T(\mathbf{y}_n) \not\rightarrow 0$. This implies that T is not continuous at the origin, but this is impossible since T is continuous. Thus we have shown that all continuous operators are bounded.

To show the converse, assume that $\|T(\mathbf{x})\| \leq n\|\mathbf{x}\|$ for some $n \in \mathbb{R}$. This implies that as $\mathbf{x}_n \rightarrow 0$, $T(\mathbf{x}_n) \rightarrow 0$. So we have that T is continuous

at the origin. Thus $\mathbf{x}_n \rightarrow \mathbf{x} \Leftrightarrow \mathbf{x}_n - \mathbf{x} \rightarrow \mathbf{0}$ implies that $T(\mathbf{x}_n - \mathbf{x}) \rightarrow \mathbf{0}$. By the linearity of T , $T(\mathbf{x}_n - \mathbf{x}) \rightarrow \mathbf{0}$ iff $T(\mathbf{x}_n) - T(\mathbf{x}) \rightarrow \mathbf{0}$. Putting all this together we get that $\mathbf{x}_n \rightarrow \mathbf{x} \Rightarrow T(\mathbf{x}_n) \rightarrow T(\mathbf{x})$. In other words T is continuous. □

Let us briefly discuss four types of operators that arise frequently: closed, adjoint, isometric and unitary. As we will soon see, certain integral transforms can be viewed as isometries between Hilbert Spaces. We will also see that unitary operators correspond to invertible transforms, and that the inverse of a unitary operator will be equal to its adjoint.

Definition 1.11. Let T be a linear transformation on a Hilbert Space \mathcal{H} . Its graph $G(T) = \{(x, Tx) \mid x \in \mathcal{H}\}$ is a subspace of $\mathcal{H} \times \mathcal{H}$ in which addition is defined by:

$$(x, Tx) + (y, Ty) = (x + y, T(x + y))$$

If $G(T)$ is closed as a subset of $\mathcal{H} \times \mathcal{H}$ then we say that the linear transformation T is *closed*.

Definition 1.12. Let $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ be a continuous linear transformation. The *adjoint* of T is the operator $T^* : \mathcal{H}_2 \rightarrow \mathcal{H}_1$ such that for all $\mathbf{u} \in \mathcal{H}_1$, and $\mathbf{v} \in \mathcal{H}_2$ we have that $(T(\mathbf{u}), \mathbf{v})_2 = (\mathbf{u}, T^*(\mathbf{v}))_1$.

Definition 1.13. Let $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ be a continuous linear transformation. T is an *isometry* if for all $\mathbf{x} \in \mathcal{H}_1$, $\|T(\mathbf{x})\|_2 = \|\mathbf{x}\|_1$.

A property of isometries that will be useful to us is that they are invertible on their images. Indeed, $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ is an isometry, implies that $\|T(\mathbf{0})\| = \|\mathbf{0}\| = 0$ and $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$. Thus, the null space of T is trivial and T is injective. Therefore, T is a bijection on its image and is invertible there.

Definition 1.14. Let $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ be a continuous linear transformation. T is *unitary* if it is a bijective isometry. Alternatively, it is an isometry such that $Im(T) = \mathcal{H}_2$.

Corollary. Unitary transformations preserve inner products. Indeed $\|T(\mathbf{x})\| = (T(\mathbf{x}), T(\mathbf{x})) = (T^*T(\mathbf{x}), \mathbf{x}) = \|\mathbf{x}\|$ implies that $T^*T = I$. Thus $T^* = T^{-1}$ and $(T(\mathbf{x}), T(\mathbf{y})) = (T^*T(\mathbf{x}), \mathbf{y}) = (\mathbf{x}, \mathbf{y})$.

2 Locally Compact Groups

While one can define many types of integral transforms, it turns out that many useful ones arise as a result of actions by locally compact groups. Furthermore many domains such as the line and n -dimensional Euclidean space are all locally compact groups, while the 2-Sphere is a locally compact set. Thus, local compactness will make its way into our discussion a number of times. In this section, we provide a brief overview of local compactness, locally compact groups and some of their properties.

Definition 2.1. A topological space is *locally compact* if every point in the space has a compact neighborhood.

It can be shown that a normed vector space is locally compact iff it is finite dimensional. This means that all of the infinite dimensional Hilbert Spaces we will see such as $\mathcal{L}^2(\mathbb{R})$ and $\mathcal{L}^2(\mathbb{S}^2)$ are not locally compact. On the other hand, the domains of these spaces of functions such as \mathbb{R} and \mathbb{S}^2 are locally compact. Another canonical example of a non-locally compact set is the set of all rational numbers \mathbb{Q} when viewed as a subset of the real line. Let us now define locally compact groups which are nothing more than topological groups that are locally compact.

Definition 2.2. A *topological group* G , is a group that is also a topological space such that the group law is a continuous map on that space as well as the map sending elements in the group to their inverses.

We will need one theorem to work with topological groups. It makes explicit the topology we have in mind when talking about locally compact groups. For a proof we refer the reader to [8].

Definition 2.3. If for each x, y in a topological space X , such that $x \neq y$ there exist open sets U, V such that $U \cap V = \emptyset$ and $x \in U, y \in V$ then X is said to be *Hausdorff*. If for each x, y in a topological space X , such that $x \neq y$ each has a neighborhood not containing the other then X is said to be T_1 .

There are topological groups that are not Hausdorff, but we will not give any examples of them and the following theorem justifies our reticence.

Theorem 2.1. If G is T_1 then G is Hausdorff. If G is not T_1 then $\overline{\{e\}}$ is a closed normal subgroup, and the quotient group $G/\overline{\{e\}}$ is Hausdorff.

In light of this theorem, it is usually sufficient to assume that a topological group is Hausdorff as a topological space. It is now simple to define a locally compact group.

Definition 2.4. A *locally compact group* is a topological group such that the underlying topology is Hausdorff and locally compact.

Theorem 2.2. Suppose H is a subgroup of a topological group G . Then if G is locally compact then so is the quotient space G/H .

Proof. Let $\psi : G \rightarrow G/H$ be the canonical quotient map. Let $U \subset G$ and $x \in G$. Let Ux denote the set $\{gx \mid g \in U\}$. Because ψ induces the quotient topology on G/H , ψ maps open sets to open sets, and also pulls open sets back to open sets. Therefore ψ is continuous and images of compact sets under continuous functions are compact.

Since G is locally compact, there is a compact neighborhood U of some $x_0 \in G$. Therefore, for any x , by the arguments above $\psi(Ux)$ is a compact neighborhood of $\psi(x) \in G/H$. Therefore, G/H is locally compact. \square

As a consequence of the above theorem we see immediately that the quotient $SO(3)/SO(2) \cong \mathbb{S}^2$ is locally compact since $SO(2)$ is a subgroup of $SO(3)$. This brings us to the next section.

3 $SO(3)$ and the 2-Sphere

This section reviews some conventions about the 2-Sphere as well as some facts about $SO(3)$ that will be referred to occasionally. We will be interested in Fourier-like expansions of functions on \mathbb{S}^2 as well as an ability to manipulate these expansions. This requires a brief treatment of spherical harmonics and their properties. Finally, we present the stereographic projection — a map from the sphere to the plane. This tool will be indispensable to us when we discuss integral transforms on the 2-Sphere.

In what follows, we parameterize the unit sphere \mathbb{S}^2 embedded in \mathbb{R}^3 using standard spherical coordinates. Thus, an element $\eta \in S^2$ will be written as:

$$\eta = (\cos(\phi) \sin(\theta), \sin(\phi) \sin(\theta), \cos(\theta))$$

with $\phi \in [0, 2\pi)$ and $\theta \in [0, \pi]$.

Let us now discuss a very important family of functions on the sphere — the spherical harmonics. In order to properly define a spherical harmonic, we first have to introduce the Legendre polynomials $P_l^m(x)$. The Legendre polynomials can be generated by the Gram-Schmidt orthogonalization process applied to the set $\{1, x, x^2, \dots, x^i, \dots\}$ in the open interval $(-1, 1)$.

Given $l \in \mathbb{N}$, $|m| \leq l$, the Legendre polynomials are defined as:

$$P_l^m(x) = \frac{(-1)^m (1-x^2)^{m/2}}{2^l l!} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l$$

We now define the spherical harmonics $Y_{lm} : \mathbb{S}^2 \rightarrow \mathbb{C}$.

$$Y_{lm}(\eta) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos(\theta)) e^{im\phi}$$

The above form for spherical harmonics is derived by viewing the Y_{lm} as everywhere regular eigenfunctions of the spherical Laplace operator. We will not present the derivation here, but it can be found in any text dealing with orthogonal functions such as [21]. The next theorem allows us to write any square-integrable function on a sphere as a series of spherical harmonics. For a proof see [4].

Theorem 3.1. The spherical harmonics constitute an orthogonal basis for $\mathcal{L}^2(\mathbb{S}^2)$.

For reasons which will become apparent in subsequent sections, we need to explain how to rotate functions on the sphere. Since, by Theorem 3.1, all square-integrable functions on the sphere can be written as linear combinations of spherical harmonics, we only need to understand how $SO(3)$, acts on spherical harmonics.

Let us briefly recount the basic properties of $SO(3)$. We list some properties without definition, but they will not be used so any reader may skip over them. The most intuitive definition for this group is as the group of all rotations of a 2-Sphere. Herein lies the connection between the two. One can also think of $SO(3)$ as the group of all orthogonal 3x3 matrices with determinant 1. $SO(3)$ is a compact, connected Lie group and is one of the classical groups. Since $SO(3)$ is isomorphic to $SU(2)/\mathbb{Z}_2$ and $SU(2)$ is homeomorphic to the 3-Sphere, $SO(3)$ can be viewed as the 3-Sphere with antipodal points identified. In our discussion, $SO(3)$ will be parameterized by Euler angles γ, β, α .

This means that each rotation in $SO(3)$ can be written as $g(\gamma, \beta, \alpha) = R_z(\gamma) \circ R_y(\beta) \circ R_z(\alpha)$. In matrix notation $R_y(\beta)$ and $R_z(\alpha)$ take the forms:

$$\begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix}, \quad \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Let $f \in \mathcal{L}^2(\mathbb{S}^2)$ and let $g \in SO(3)$ be a rotation. We define the rotation operator $\Lambda(g) : \mathcal{L}^2(\mathbb{S}^2) \rightarrow \mathcal{L}^2(\mathbb{S}^2)$ as:

$$[\Lambda(g)f](\eta) = f(g^{-1}\eta). \quad (1)$$

An important property of spherical harmonics is that the subspaces

$$\mathcal{L}^2(\mathbb{S}^2) \supset Y_l = \{Y_{lm} \mid |m| \leq l\}, \quad l \in \mathbb{N}$$

are invariant under the action of $SO(3)$. It is known that as a result of this [5], a rotated spherical harmonic function can be expressed by:

$$\Lambda(g)Y_{lm}(\eta) = \sum_{|n| \leq l} U_{mn}^l(g)Y_{ln}(\eta)$$

where the $(2l+1) \times (2l+1)$ matrix $U_{mn}^l(g)$ is given by

$$U_{mn}^l(g(\gamma, \beta, \alpha)) = e^{-im\gamma} P_{mn}^l(\cos(\beta)) e^{-in\alpha}$$

Finally, we note that:

$$U_{mn}^l(g_2 g_1) = \sum_{|k| \leq l} U_{mk}^l(g_2) U_{kn}^l(g_1)$$

3.1 Stereographic Projection

We now present a tool we will use to help us construct an integral transform on the sphere — the stereographic projection. We will also use it to help us define useful functions and filters by lifting them via the inverse stereographic projection from the plane to the sphere.

The stereographic projection is a way to map points on the sphere to the plane. Specifically, the stereographic projection Π is a map $\Pi : S^2 \setminus \{(0, 0, -1)\} \rightarrow \mathbb{R}^2$. Its action is depicted in figure 1.

Points from the sphere are mapped to the plane tangent to the sphere at its north pole. The geometric interpretation for the projection is as follows. Given a point ω on the surface of the sphere, we draw a ray from the south pole of the sphere going through ω . This ray intersects the plane at some point z . Then the stereographic projection maps the point ω to z . We can write this down explicitly:

$$x(\phi, \theta) = \frac{2 \sin(\theta) \cos(\phi)}{1 + \cos(\theta)} \quad y(\phi, \theta) = \frac{2 \sin(\theta) \sin(\phi)}{1 + \cos(\theta)}$$

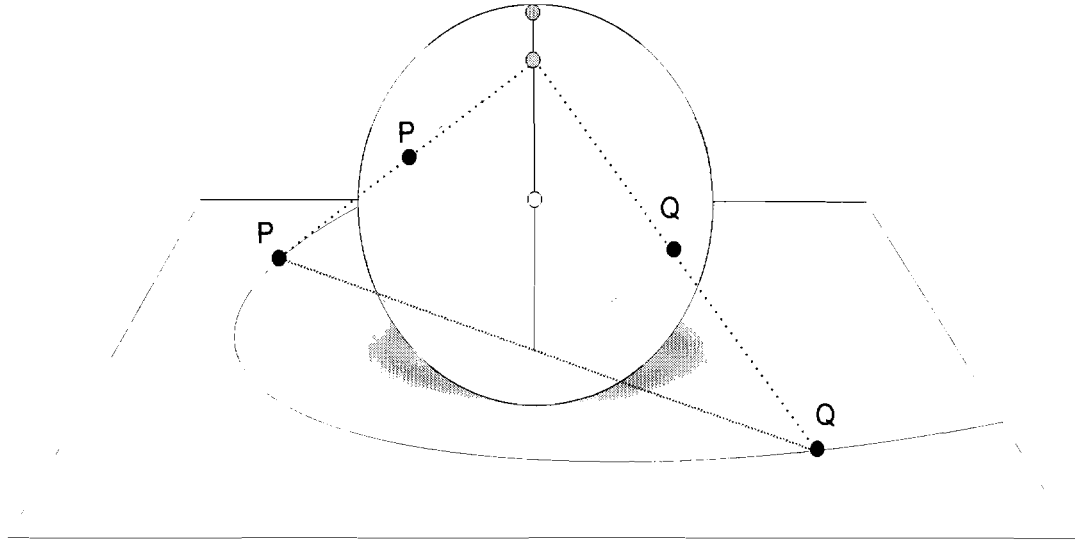


Figure 1: The stereographic projection with two points and their images labelled on the sphere and the plane. Courtesy Christopher Geyer.

The stereographic projection is invertible. The explicit form of the inverse $\Pi^{-1} : R^2 \rightarrow S^2 \setminus \{(0, 0, -1)\}$ is given by:

$$\phi(x, y) = 2 \tan^{-1}\left(\frac{\sqrt{x^2 + y^2}}{2}\right) \quad \theta(x, y) = \tan^{-1}\left(\frac{y}{x}\right)$$

Basic Measure Theory

1 On the use of Measure Theory

In this section we take a lightning fast tour of measure theory. The aim is to understand how to integrate functions over abstract sets. This is the application that is the most clear motivation for the development of measure theory. There are many functions one can define, even on the real line, that the Riemann integral cannot handle. Measure-theoretic ideas allow for the development of the more general Lebesgue integral which resolves this problem. To be sure, there still exist sets that are not well-behaved even in a measure-theoretic sense. However, their constructions are frequently non-trivial and they rarely arise naturally in practice.

In particular, we will be interested in how to integrate functions over locally compact groups. Applications of, what at first glance seems an unnecessarily abstract topic, are numerous. A classical example that will be described in a subsequent chapter is that of convolution on the group of rotations of the 2-Sphere $SO(3)$. There are many applications of this in physics and computer vision, some of which we will explore. From a signal processing standpoint, measure theory will allow us to analyze continuous signals that arise on many types of spaces. Continuous integral transforms, possibly the most important tools used in signal processing, need to be defined on a variety of domains such as surfaces that can be embedded into Euclidean space. Furthermore, the development of continuous transforms almost always leads the way to discrete transforms on the same domains opening the way to even more applications.

We will proceed in the following way. First, we will define a large class of functions, measurable functions, that the theory will be able to handle. In a manner that is analogous to topology, spaces that carry measurable functions will be emphasized. We then introduce the concept of a measure as a basic tool that can tell us something meaningful about the domains on which measurable functions are defined. The measure will tell us over which parts of the domain we need to pay attention to what values the function takes on and over which parts we do not. This immediately gives rise to the Lebesgue integral. We will then shift our focus to defining an integral with respect to an appropriate measure, Haar measure, on locally compact groups.

Our treatment will be rapid, but we will touch on all of the relevant points. This, however, will not do any justice to the field so the interested

reader should consult standard references on the subject [12] and [20]. Much of the material on Haar measure can also be found in [13] and [8].

2 Measure Spaces

In measure theory, integration is performed on something called measure spaces and in this section, it is our goal to define them. To do this, measurable spaces first have to be defined. A measure space is a measurable space with a positive function, called a measure, defined on it. We present all the necessary definitions and provide examples.

Definition 2.1. A class of subsets \mathfrak{M} of a set X is a σ -algebra in X if it has the following properties:

1. $X \in \mathfrak{M}$.
2. \mathfrak{M} is closed under complements.
3. \mathfrak{M} is closed under countable unions.

Definition 2.2. Let X be a set and \mathfrak{M} be a σ -algebra on X . Then (X, \mathfrak{M}) is called a *measurable space* and members of \mathfrak{M} are called *measurable sets*. Frequently a measurable space (X, \mathfrak{M}) is denoted by X when the exact nature σ -algebra is unimportant.

We now define precisely those functions that we will be able to work with.

Definition 2.3. Let X be a measurable space and Y be a topological space. Let $f : X \rightarrow Y$ be a map. If for all open sets $A \subseteq Y$, $f^{-1}(A)$ is measurable, then we say f is a *Borel measurable function*.

We include a number of theorems without proof. They are not difficult to prove, but can be found in any text dealing with elementary measure theory.

Theorem 2.1. Let A be a measurable subset in X . Define the function χ_A by:

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A, \end{cases}$$

χ_A is called the *characteristic function* of the set A .

Theorem 2.2. Let F be a family of subsets of a set X . There exists a smallest σ -algebra \mathfrak{M}_s in X so that $F \subset \mathfrak{M}_s$.

An important and useful example of a measurable space is the following construction. Let X, Y be topological spaces. Let \mathcal{B} be the smallest σ -algebra containing all of the open sets of X . Its existence is guaranteed by the last theorem. Let $f : X \rightarrow Y$ be a continuous function. Then it immediately follows that it is also measurable with respect to \mathcal{B} . It is easy to see that these functions are *Borel measurable* and, therefore, we call the members of \mathcal{B} , *Borel sets*.

Now we turn to the last important piece of the abstract theory – measures. These functions defined on σ -algebras are the needed tools to properly define measure spaces.

Definition 2.4. Let X be a measurable space and \mathfrak{M} be the associated σ -algebra. A *measure* is a function $\mu : \mathfrak{M} \rightarrow [0, \infty]$ that satisfies the following properties:

1. $\exists A \in \mathfrak{M}$ such that $\mu(A) < \infty$.
2. If $\{A_i\}$ is a disjoint countable family of subsets of \mathfrak{M} then

$$\mu \left(\bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} \mu(A_i)$$

Finally, we present the definition of a measure space and give some simple examples. More examples, will be constructed in subsequent sections.

Definition 2.5. A *measure space* is a measurable space which has a measure defined on the σ -algebra of its measurable sets.

Example 2.1. Let X be any set and the σ -algebra be the powerset of X . For every $A \in \mathfrak{M}$ define the measure $\mu(A)$ as follows:

$$\mu(x) = \begin{cases} |A| & \text{if } A \text{ is finite,} \\ \infty & \text{if } A \text{ is infinite,} \end{cases}$$

This is called a *counting measure* on the set X .

Example 2.2. Let X be any set and the σ -algebra be the powerset of X . Any measure with the property that $\mu(X) = 1$ is called a *probability measure*.

We end this section with a comment about the sets which have measure zero. When using measure theoretic constructions, one typically wants to construct the underlying measure space in such a way that the important properties, for whatever the application is, are true everywhere. Of course, this is not always possible and the next best thing is to construct the measure space in such a way that the sets on which the desired property fails have measure 0. Then, we say that the property holds *almost everywhere*.

3 Abstract Integration

In this section we extend the notion of integration to measurable functions. The idea is to first describe a small set of functions and explain how to integrate them. We will then see how to approximate every function (and its integral) by a sequence of members of this set (and their integrals). Therefore, as part of this approach, we will have to explain what is meant by convergence in measure. For brevity, we omit most of the proofs.

Definition 3.1. Let s be a function on a measure space X that takes on a finite number of non-negative values $\{\alpha_1, \dots, \alpha_n\}$ in its co-domain. Then s is called a *simple function*. Let $E_i = \{x \mid s(x) = \alpha_i\}$, then every simple function can be written as:

$$s = \sum_{i=1}^n \alpha_i \chi_{E_i}(x)$$

where χ_{E_i} is the characteristic function of E_i .

Observe that the definition implies that s is measurable only if A_i is measurable for all $1 \leq i \leq n$. The definition of the integral of a simple function is now very natural to state.

Definition 3.2. Let s be a simple function on a measure space X and let $E \subseteq X$. We define the integral of s as:

$$\int_E s d\mu = \sum_{i=1}^n \alpha_i \mu(A_i \cap E)$$

In analysis, there are a number of different types of convergence such as uniform convergence and absolute convergence. Furthermore, one can attribute “convergence-like” properties to sequences. For instance one can say that a sequence is Cauchy. Likewise, on a measure space it is possible

to define many types of convergence and characterize sequences in a similar fashion. We next define mean fundamental sequences and convergence in measure.

Definition 3.3. A sequence $\{f_n\}$ of integrable functions is *mean fundamental* if

$$\int |f_n - f_m| d\mu \rightarrow 0 \text{ as } n, m \rightarrow \infty$$

Definition 3.4. A sequence of measurable functions $\{f_n\}$ *converges in measure* to a measurable function f if, $\forall \epsilon > 0, \lim_{n \rightarrow \infty} \mu(\{x : |f_n(x) - f(x)| \geq \epsilon\}) = 0$.

We are now ready to discuss integrability for any measurable function.

Definition 3.5. Let X be a measure space and let f be a measurable function on X . We say that f is integrable if there exists a mean fundamental sequence of simple functions $\{f_n\}$ that converges in measure to f .

It can be shown that for any measurable function f there exists a mean fundamental sequence of simple functions that converge in measure to f . Once again, the integral of a measurable function f can be defined naturally by:

$$\int_X f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu$$

Finally the integral of a measurable function over a subset E of X is given by:

$$\int_E f d\mu = \int \chi_E f d\mu$$

4 Haar Measure

In this section we discuss measures and integration on locally compact groups. While the theory presented in the previous two sections is general enough to include locally compact groups, it remains to discuss how to construct measures on them explicitly. Furthermore, we would like the measures on groups to have an additional property of invariance under group actions. We will supply a motivation for this additional requirement following its definition.

Definition 4.1. Let G be a group and S a set. If to every $g \in G$ we can associate a transformation of S into itself: $T_g : S \rightarrow S$ such that for all $g, h \in G$, $T_{gh} = T_g T_h$ then we say that G acts on S on the left. Each T_g is called a *left action*.

On the other hand, if for all $g, h \in G$, $T_{gh} = T_h T_g$ then we say that G acts on S on the right. Each T_g is called a *right action*.

A bit of notation. Let G be a group acting on a set S on the left. If $I \subseteq S$ and T_g a left action then we define $T_g(I)$ to be $\{y | y = gx, \text{ for some } x \in I\}$.

Example 4.1. Let $G = S = \mathbb{R}$. For every $g \in G$ let $T_g : \mathbb{R} \rightarrow \mathbb{R}$ be defined as $T_g(x) \rightarrow x + g$. In this case because G is an abelian group we do not make distinctions between left and right actions.

We will use the above example to motivate why we might want the measure on any set to be invariant under left or right actions. While this example simplifies things greatly, it still has enough motivating power. From chapter 2 we know that \mathbb{R} is a locally compact group. Suppose we define a measure μ on it that is not left or right invariant. Now consider the closed interval $[a, b]$ on the real line. Under the action by T_x the interval is shifted to $[a+x, b+x]$. It would not be very “natural” if $\mu([a, b]) \neq \mu([a+x, b+x])$. Therefore, we seek measures on sets such that they are invariant under left or right group actions.

Definition 4.2. Let G be a locally compact group, S a set, $I \subseteq S$ and μ a measure on S . We say that μ is *left-invariant* if for all left actions T_g of G on S , $\mu(T_g I) = \mu(I)$. Right invariance is defined similarly. A left-invariant measure is called a *left Haar measure*. A right-invariant measure is called a *right Haar measure*.

In the above definition we assume that in order to define μ , a σ -algebra has been made explicit. An obvious question at this point is: “How do we know that there is always a Haar measure for a set S with respect to action by a locally compact group G ?” In 1940, Weil answered this question positively for the case when $G = S$. This is the reason, that we mentioned constructing measures for locally compact groups. Keep in mind that this result does not imply that when $G \neq S$ there cannot be a Haar measure on S . A proof can be found in [13].

Theorem 4.1. Let S be a set and G a locally compact group. If $G = S$ is a locally compact group and let the left action be defined by multiplication on the left. Then there exists a unique left and right invariant Haar measure on S .

Before presenting an example let us briefly touch upon the difference between left and right invariance. It is easy to see that for all commutative groups left and right Haar measure is one and the same. All groups that have this property are called *unimodular* groups. Nevertheless, there are examples of groups that are non-commutative and for which left Haar measure is the same as the right Haar measure. For instance, all finite groups with counting measure are unimodular. Let us give an example of a Haar measure.

Example 4.2. Let $G = \langle \mathbb{R} \setminus \{0\}, \cdot \rangle$. Then the Haar measure on this multiplicative group is $dx/|x|$. To verify this, recall the one dimensional change-of-variable formula. Let $y(x)$ be an injective differentiable mapping with compact support of an open set $E \subset \mathbb{R}^n$ into \mathbb{R}^n such that $\frac{dy}{dx} \neq 0$. Then:

$$\int_G f(y)dy = \int_G f(y(x))\left|\frac{dy}{dx}\right|dx$$

Now consider how G acts on itself. Fix $a \in G$ and let $\tau : G \rightarrow G$ be the corresponding left action: $x \rightarrow ax$. Then by applying the change of variable formula we have:

$$\begin{aligned} |a| \int_G f(\tau(x)) \frac{dx}{|x|} &= \int_G f(\tau(x)) \left| \frac{d\tau(x)}{dx} \right| \frac{dx}{|x|} = \\ &= \int_G f(x) \frac{dx}{|\tau^{-1}(x)|} = \int_G f(x) \frac{dx}{|x/a|} = |a| \int_G f(x) \frac{dx}{|x|} \end{aligned}$$

Therefore, we have that $\frac{dx}{|x|}$ is a left Haar measure on G since by dividing through by $|a|$ we have:

$$\int_G f(\tau(x)) \frac{dx}{|x|} = \int_G f(x) \frac{dx}{|x|}$$

It turns out that the above construction is a particular case of a more general theorem that can be used to derive Haar measures for a large family of locally compact groups. We conclude this chapter by stating this theorem. A proof is omitted because it is a straightforward generalization of the example one given above.

Theorem 4.2. Suppose the underlying topological space of G is an open subset of \mathbb{R}^n and left translations are given by affine maps: $(x, y) \rightarrow A(x)y + b(x)$, where $A(x)$ is an invertible linear transformation on \mathbb{R}^n and $b(x)$ is a vector in \mathbb{R}^n . Then $|\det A(x)|^{-1}dx$ is a left Haar measure on G , where $dx = dx_1 \dots dx_n$ is the usual measure on \mathbb{R}^n . (Adapted from [8])

Basic Representation Theory

1 On the use of representations

One of the goals of this thesis is to explain how square-integrable representations of groups give rise to integral transforms. Thus, the main purpose of this chapter should be to develop representation theory to a level that is sufficient for our needs. However, this is only part of the program and the most important goal is to point out that groups, not group representations, are the objects that give rise to integral transforms.

In the broadest sense, representation theory is a study of groups. Therefore, we are always more concerned with how particular representations reflect the characteristics and structure of the underlying groups, rather than with properties of the representations themselves. The general method of representation theory is to study groups through understanding how they act on various vector spaces. It turns out that these actions are severely restricted by the structure of the group, thus giving us interesting insights. As we will see in the next chapter, actions of certain groups can give rise to integral transforms. So while representations will be used to encode these actions, it is crucial to understand that it is the group that gives rise to the transforms and everything else is nothing but machinery. Let us now turn to the subject matter.

2 Definition of a Representation

We structure our presentation based on the observation, made by many authors, that the theory of representations of locally compact groups is very similar to the theory of representations of finite groups. Therefore, we start with the latter and via the addition of measure-theoretic concepts such as Haar measure, we generalize the discussion to the former. For more in-depth treatment of the various material presented, consult [22] and [9].

Let \mathcal{H} be a Hilbert space on a field F . The group of non-singular linear transformations from \mathcal{H} to itself will be denoted by $GL(\mathcal{H})$. We now give the main definition.

Definition 2.1. A *linear representation* of a group G is a homomorphism R from G to $GL(\mathcal{H})$. The dimension of \mathcal{H} is called the *degree* of the representation.

It is possible to define representations in a slightly more general way by

studying homomorphisms into groups of invertible linear maps over any vector space. Indeed, our use of Hilbert Spaces is a bit of an overkill since their geometrical structure is not required. Nevertheless, the only examples we consider will have a Hilbert Space as the underlying vector space. Therefore, this restriction makes the discussion more focused on what is important.

In general, it is preferable to work without fixing an explicit basis for \mathcal{H} . When the degree of the representation is finite, we can always construct a matrix representation for some linear representation R , by picking an explicit basis for \mathcal{H} . The choice of this basis defines an isomorphism $\phi : GL(\mathcal{H}) \rightarrow GL_n(F)$ by $\phi(T) \rightarrow \text{matrix of } T$. Thus, we come upon the following definition:

Definition 2.2. Let $n \in \mathbb{Z}^+$. An n -dimensional *matrix representation* of a group G is a homomorphism from G to $GL_n(F)$.

We will use matrix representations only for purposes of explicit computation and try to phrase most of the general results in the language of linear representations. Let us end this section with two examples of representations.

Example 2.1. Let $G = \langle \mathbb{Z}_2, +_2 \rangle$ and $v \in \mathcal{H}$. The first representation ϕ , defined by $\phi(0)v = v$ and $\phi(1)v = v$ is called the trivial representation. The second representation ψ defined by $\psi(0)v = v$ and $\psi(1)v = -v$ is the only non-trivial representation of $\langle \mathbb{Z}_2, +_2 \rangle$.

Example 2.2. Groups can have many different representations, especially when one varies the underlying vector space. Consider $G = S_3$ and $v \in \mathcal{H}$. Let ϕ be a one-dimensional representation of G on \mathbb{R} defined by: $\phi(\sigma)v = \text{sgn}(\sigma)v$. On the other hand we can consider a two-dimensional representation of S_3 on $GL_2(\mathbb{R})$. Here S_3 is given by the following six matrices:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ -1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$$

3 Representations of Finite Groups

As previously mentioned, the theory of representations of locally compact groups is very close to the theory of representations of finite groups. We start with an intriguing theorem about the latter.

Theorem 3.1. Let $R : G \rightarrow GL(\mathcal{H})$ be a representation of a finite group G . Then \mathcal{H} has an inner product (\cdot, \cdot) in which each $R(g)$ is unitary. Thus, for all $v, w \in \mathcal{H}$ and for all $g \in G$

$$(R(g)v, R(g)w) = (v, w)$$

Proof. Let $(\cdot, \cdot)_0$ be an inner product on V . We define a new inner product (\cdot, \cdot) by:

$$(v, w) = \frac{1}{|G|} \sum_{g \in G} (R(g)v, R(g)w)_0$$

We leave the verification that the above is an inner product to the reader and proceed to show that each $R(g)$ is unitary. For a fixed $h \in G$, it is clear that the map $g \rightarrow gh$ is a bijection as g runs through G .

$$\begin{aligned} (R(g)v, R(g)w) &= \frac{1}{|G|} \sum_{g \in G} (R(g)R(h)v, R(g)R(h)w)_0 \\ &= \frac{1}{|G|} \sum_{g \in G} (R(gh)v, R(gh)w)_0 = \frac{1}{|G|} \sum_{g \in G} (R(g)v, R(g)w)_0 = (v, w) \end{aligned}$$

□

The following corollary is beautiful enough to deserve mention. It shows the well-known result that all eigenvalues of unitary operators are roots of unity. Nonetheless, it gives a good taste for how representation theory applies group theory to the study of linear spaces.

Corollary. All $R(g)$ have eigenvalues λ with absolute value 1. The proof is trivial. Since G is a finite group, $\exists n \in \mathbb{Z}^+$ such that $g^n = e$. Thus $R(g)^n = I$, therefore $\lambda^n = 1$ implying that λ is a root of unity.

In any case, the importance of the above theorem is that, in the finite case, we can now restrict our attention to unitary representations of a group. In fact, most authors consider only unitary representations and define a representation R of a group G as a homomorphism $R : G \rightarrow U(\mathcal{H})$ where $U(\mathcal{H})$ is the space of unitary transformations acting on \mathcal{H} . We will adopt this definition as well.

Definition 3.1. A *unitary representation* of a group G is a homomorphism R from G into $U(\mathcal{H})$.

In this setting, it is natural to define the notion of equivalence of representations.

Definition 3.2. Let $R : G \rightarrow U(\mathcal{H}_1)$ and $R' : G \rightarrow U(\mathcal{H}_2)$ be two unitary representations of the same finite group. We say that R and R' are *unitarily equivalent* if there exists a unitary transformation $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ such that for all $g \in G$:

$$R'(g) = TR(g)T^{-1}$$

To better understand the goals of representation theory as well as harmonic analysis (i.e. which group representations does one study and why?) we need to introduce the concepts of direct sum representations and irreducibility. Notice the slight abuse of notation for pairs of elements vs. inner products. However, the notation is consistent in that arguments for inner products are in bold, while elements of pairs are not.

Definition 3.3.

- a. If $\mathcal{H}_1, \mathcal{H}_2$ are Hilbert Spaces, then so is $\mathcal{H}_1 \oplus \mathcal{H}_2$ with inner product $((\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2)) = (\mathbf{x}_1, \mathbf{x}_2) + (\mathbf{y}_1, \mathbf{y}_2)$.
- b. If $A, B \in GL(\mathcal{H})$ then the direct sum operator $A \oplus B$ on $\mathcal{H}_1 \oplus \mathcal{H}_2$ is defined by

$$(A \oplus B)(\mathbf{x}, \mathbf{y}) = (A\mathbf{x}, B\mathbf{y})$$

From the above definitions we see that if $A \in U(\mathcal{H}_1)$, $B \in U(\mathcal{H}_2)$ then $A \oplus B \in U(\mathcal{H}_1 \oplus \mathcal{H}_2)$. Therefore, if R and R' are representations of G on $U(\mathcal{H}_1), U(\mathcal{H}_2)$ respectively, then $U \oplus V$ defined by

$$(R \oplus R')(g) = R(g) \oplus R'(g)$$

is also a representation of G . It is called the *direct sum representation*.

Definition 3.4. R is an *irreducible representation* if it cannot be written as a direct sum of non-trivial representations.

Theorem 3.2. Any finite-dimensional linear representation can be written as a direct sum of irreducible representations.

Proof. We proceed by induction on the degree of R . If $\deg(R) = 1$, we are done since all one-dimensional representations are irreducible. Suppose we have the result for all representations of degree $< d$. Let $\deg(R) = d$, if R is irreducible we are done. Otherwise, $R = R' \oplus R''$, with $\deg(R') < d$ and $\deg(R'') < d$. Thus, by induction, the theorem follows. \square

Taking into consideration the definitions above, two questions arise. For a given finite group G , what are all of its irreducible unitary representations? Can we construct all possible representations for G using direct sums of irreducible representations? These are some of the questions that representation theory tries to answer. Moreover, the general theory is not only concerned with finite groups. Many others such as semi-simple Lie groups, locally compact groups and non-locally compact non-abelian groups are studied. In the next section, we will generalize the definitions and theorems given in this section to locally compact groups since these are the ones we are concerned with.

4 Representations of Locally Compact Groups

Most of the results that we state in this section are analogues of the finite-dimensional case. It should also be noted that now that the group is a topological group, there are many details suddenly arise when talking about representations. For instance, when we will speak about continuous representations, one has to specify which topology is being discussed. One can also ask questions about whether or not certain assumptions in the definitions are necessary or if they are implied by weaker assumptions. We sweep all of these details aside to get the results we need. Thus, if certain definitions and results seem unnatural or contrived, it is for the sake of clarity.

We assume that for any locally compact group G a left Haar measure μ is given. The first order of business, is to generalize theorem 3.1.

Theorem 4.1. Let G be a locally compact group. Let $\phi : G \rightarrow GL(\mathcal{H})$ be a homomorphism of G onto a set of linear maps on a separable complex Hilbert Space \mathcal{H} . Suppose ϕ is measurable and bounded and that for all $v \in \mathcal{H}$, $\phi(g)v$ is not almost everywhere 0 w.r.t to the appropriate measure on G . Then, there exists an inner product on \mathcal{H} on which all $\phi(g)$ are unitary.

Proof. This is very similar to the finite case. Let $(\cdot, \cdot)_0$ be an inner product on \mathcal{H} . We define a new inner product (\cdot, \cdot) by:

$$(v, w) = \int_{g \in G} (\phi(g)v, \phi(g)w)_0 d\mu(g)$$

Since \mathcal{H} is separable, let $\{e_i\}_{i=1}^{\infty}$ be a countable orthonormal basis. We have to argue that the newly defined inner product is integrable and strictly

positive. This is easy to see because

$$(\mathbf{v}, \mathbf{w}) = \sum_{j=1}^{\infty} (\phi(g)\mathbf{v}, e_j)_0 (e_j, \phi(g)\mathbf{w})_0$$

Since for all $\mathbf{v} \in \mathcal{H}$, $\phi(g)\mathbf{v}$ is not almost everywhere 0, the function is strictly positive. It is also easy to verify that each $\phi(g)$ leaves this inner product invariant. \square

We can now define what is traditionally understood to be a representation of a locally compact group.

Definition 4.1. Let G be a locally compact group. Let R be a continuous homomorphism $G \rightarrow U(\mathcal{H})$, to the space of unitary transformations on a separable complex Hilbert Space \mathcal{H} . Then we say that R is a *unitary representation* of G in \mathcal{H} .

For brevity, we only give one example of such a representation and defer the rest to the next chapter, when we work them out explicitly for two groups.

Example 4.1. Let $G = \langle \mathbb{R}, + \rangle$ and $\mathcal{H} = \mathcal{L}^2(\mathbb{R})$. For every $k \in \mathbb{R}$ we associate a translation operator $T_k : \mathcal{L}^2(\mathbb{R}) \rightarrow \mathcal{L}^2(\mathbb{R})$ whose action on any $f(x) \in \mathcal{L}^2(\mathbb{R})$ is defined by $T_k(f(x)) = f(x+k)$. It is easily seen that the map $k \rightarrow T_k$ is an infinite dimensional, continuous and unitary representation of G in \mathcal{H} .

Let us now generalize some of the definitions and a theorem from the previous section.

Definition 4.2. Two representations of locally compact groups R and R' on spaces \mathcal{H}_1 and \mathcal{H}_2 are called *unitarily equivalent* if there is a unitary transformation $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ so that for all $g \in G$, $R'(g) = TR(g)T^{-1}$.

Definition 4.3. A representation R of a locally compact group G is called *irreducible* if it cannot be written as a direct sum of non-trivial representations.

Example 4.2. Let $G = \mathbb{R}$ and $\mathcal{H} = \mathcal{L}^2(\mathbb{R})$. For every $k \in \mathbb{R}$ we associate a modulation operator $T_k : \mathcal{L}^2(\mathbb{R}) \rightarrow \mathcal{L}^2(\mathbb{R})$ whose action on any $f(x) \in \mathcal{L}^2(\mathbb{R})$ is defined by $T_k(f(x)) = e^{ikx}f(x)$. It is easily seen that the map $k \rightarrow T_k$ is an infinite dimensional, continuous and unitary representation of G in \mathcal{H} .

Theorem 4.2. Every representation of a locally compact group G is equivalent to a direct sum of irreducible representations.

We need to mention another extremely important property of irreducible representations that will be used in the next chapter. It will be used in proving the theorem that the described integral transforms are isometries.

Theorem 4.3. Let G be a locally compact group, R a continuous, unitary, irreducible representation of G in $GL(\mathcal{H})$ and $\mathbf{h} \in \mathcal{H}$ a vector. Then the span of the set $\{R(g)\mathbf{h} \mid g \in G\}$ is dense in \mathcal{H} .

Proof. Assume to the contrary. Then there exists a non-zero vector $\mathbf{u} \in \mathcal{H}$ such that for all $g \in G$, $(R(g)\mathbf{h}, \mathbf{u}) = 0$. Let $V = \text{span}\{\mathbf{u}\}$ be a subspace of \mathcal{H} . It is invariant under action by R since by the unitarity of R :

$$(R(g_1)\mathbf{h}, R(g_2)\mathbf{u}) = (R(g_2^{-1}g_1)\mathbf{h}, \mathbf{u}) = 0$$

If V is invariant under action by R , it is easy to verify that so is V^\perp . Then in the block decomposition of $GL(\mathcal{H})$, we have that:

$$R(g) = \begin{pmatrix} R_1(g) & 0 \\ 0 & R_2(g) \end{pmatrix}$$

Thus $R = R_1 \oplus R_2$, a contradiction since R is irreducible. \square

We can now give an example of a reducible representation and we will use the above theorem to show that it is reducible.

Example 4.3. Let $G = \mathbb{R}$ and $\mathcal{H} = \mathcal{L}^2(\mathbb{R}, d\mu)$. The representation R associating a translation operator T_k to every $k \in \mathbb{R}$ such that $[T_k f](x) = f(x + k)$ is reducible. Consider a set of square-integrable functions indexed by the set I : $\{f_i\}$. Further, let us denote the set of Fourier Transforms $\{\mathcal{F}(f_i)\}$ of functions in this set by S .

We will show that the set S is invariant under action by R . The representation R acts by translation on square-integrable functions in the time-domain. Therefore, it acts by modulation on square-integrable functions in the frequency domain. But the set S is closed under modulation. By an application of the previous theorem this implies that R is a reducible representation.

It turns out that irreducibility is a sufficient but not necessary restriction on a representation for the above theorem to hold. We will see examples in chapter 6 of representations that satisfy this property but are not irreducible. In general, the representations that satisfy the above theorem are called *cyclic*.

Integral Transforms

In this chapter we come to the central topic of this paper – integral transforms. Using some of the tools developed in the previous chapters we will show how one can associate an integral transform with a group acting on a Hilbert Space. This would be a mere curiosity if not for the abundance of specific instances of this construction that are very well-known and frequently used in a variety of applications. In particular, the Continuous Wavelet Transform (CWT) and the Windowed Fourier Transform (WFT) arise in this context.

This chapter is subdivided into two sections. The first deals with the general construction described above. Nevertheless, it turns out that for certain pairs of groups and Hilbert spaces the general construction is impossible because key properties necessary for defining the transform do not hold given the nature of the particular group and Hilbert Space. For instance, it is impossible to develop the Continuous Wavelet Transform on the n -Sphere along the lines that will be described in the next section. Therefore, we will devote the second section of the chapter to describing methods that can be used to extend the general framework to such scenarios.

To make the task of absorbing all of this information easier, examples of the general theory will be handled in the next chapter, while the focus in this one will be solely on developing the necessary theory. This organization is also useful because it establishes the concept of the *integral transform* as an abstract entity. Indeed, just as abstract group theory can be used to study a large number of groups, a general theory for integral transforms will make it easier to single out important properties of many different transforms across the board. The approach is to view each transform as a transformation between two spaces and investigate whether or not this map has properties such as boundedness, continuity and unitarity.

In reading through this general framework, the reader should keep the following in mind. Whenever one deals with an integral transform, such as the familiar Fourier Transform, one of the most important properties is invertibility. Imagine applying a Fourier Transform to some function and getting a set of coefficients. If one could not recover the original function (or some approximation of it) from these coefficients, the number of applications of the Fourier Transform would be drastically reduced. Thus, once we have finished presenting the basic definitions and construction for the integral transform, we will focus on proving that it is invertible. In order to do this, we first show that the transform is an isometry and then explain how to

invert it on its range.

1 General Construction

The discussion in this section is based on [10], [1] and [14]. First, let us set down some notation. In the remainder of this chapter, let \mathcal{H} be a Hilbert Space, G be a locally compact group and μ a left Haar measure on G . Recall that the existence of μ is guaranteed by theorem 4.1. We begin the construction by letting G act on a Hilbert Space \mathcal{H} . Specifically, the action of G on \mathcal{H} will be determined by a representation of G in the space of linear transformations on \mathcal{H} . Let $R : G \rightarrow U(\mathcal{H})$ be a continuous unitary representation of G in the space of invertible linear transformations on \mathcal{H} .

In many cases, one can choose from a number of unitarily inequivalent continuous representations and so we will have to describe in each instance which representation we are using. In addition to the above properties, the representation also has to be square-integrable. Let us define what this means.

Definition 1.1. A non-zero vector $\mathbf{h} \in \mathcal{H}$ is called *admissible* if:

$$\int_{g \in G} |(R(g)\mathbf{h}, \mathbf{h})|^2 d\mu < +\infty \quad (2)$$

Notice that the integral is taken over the group G and not over the domain of signals in \mathcal{H} . The inner product is being treated as a function over G since it depends on the representation R which in turn is a function on G . Also, we need to mention that admissibility is a concept that will come up again when we speak about the Continuous Wavelet Transforms in the next chapter. It is an important property in its own right and is intimately related to the admissibility condition that is well-known in signal processing and wavelet literature.

Definition 1.2. We say that R is a *square-integrable representation* if R is irreducible and if there exists at least one admissible vector $\mathbf{h} \in \mathcal{H}$.

Admissibility is a necessary condition because it does not follow from irreducibility. There are irreducible representations that are not square-integrable, for instance the one in the following example.

Example 1.1. Let $\mathcal{H} = \mathcal{L}^2(\mathbb{S}^1)$. $R : \mathbb{R} \rightarrow \mathbb{S}^1$ be a continuous irreducible representation such that $R : x \rightarrow e^{iax}$. Consider the integral for all possible

values of b :

$$\begin{aligned} \int_{x \in \mathbb{R}} |(R(x)e^{imx}, e^{ibx})|^2 dx &= \int_{x \in \mathbb{R}} \left| \int_{\mathbb{R}} e^{imx} e^{ibx} e^{ibx} \right|^2 dx \\ &= \int_{x \in \mathbb{R}} \left| \int_{\mathbb{R}} e^{i(a+b+m)x} \right|^2 dx = +\infty \end{aligned}$$

Since R is a surjective homomorphism, there are no admissible vectors for which the above integral is finite and therefore R is not a square-integrable representation.

We are now ready to define the integral transform that we have waited for.

Definition 1.3. Let R be a square-integrable representation and let $\mathbf{h} \in \mathcal{H}$ be a nonzero admissible vector. Associate to \mathbf{h} the positive number c_h defined by:

$$c_h = \frac{1}{\|\mathbf{h}\|^2} \int_{g \in G} |(R(g)\mathbf{h}, \mathbf{h})|^2 d\mu$$

Then for all $g \in G$ and any $\mathbf{f} \in \mathcal{H}$ consider the complex-valued function $L_h \mathbf{f}$ defined by:

$$(L_h \mathbf{f})(g) = \frac{1}{\sqrt{c_h}} (R(g)\mathbf{h}, \mathbf{f})$$

We refer to $L_h \mathbf{f}$ the *left integral transform of \mathbf{f}* and to \mathbf{h} as the *analyzing vector*.

Once again observe that the left integral transform of \mathbf{f} is a function defined on the group G . This definition clearly shows why the choice of the group G is the single most important part of the construction. Also notice that if \mathbf{h} were not an admissible vector, then c_h would be infinite and $L_h \mathbf{f}$ would not be well-defined. However, this is not the only role that admissibility serves. It is also used in the proof which shows that the map $\mathbf{f} \rightarrow L_h \mathbf{f}$ is an isometry from \mathcal{H} to $\mathcal{L}^2(G, d\mu)$.

Let us now discuss some properties of this transform. In particular, we will try to point out why the various restrictions placed on the representation R such as continuity and irreducibility are necessary for these properties to hold. We will first show that $L_h \mathbf{f}$ is a continuous function of G . To do this we will use the fact that R is a continuous representation.

Theorem 1.1. For any $\mathbf{f} \in \mathcal{H}$, the left integral transform of \mathbf{f} is a continuous function of G .

Proof. Fix a vector $\mathbf{y} \in \mathcal{H}$. By the bilinearity of the inner product and the Cauchy-Schwartz inequality we have that:

$$\begin{aligned} |(\mathbf{x}_1, \mathbf{y}) - (\mathbf{x}_2, \mathbf{y})| &= |(\mathbf{x}_1 - \mathbf{x}_2, \mathbf{y})| \\ &\leq \sqrt{(\mathbf{x}_1 - \mathbf{x}_2, \mathbf{x}_1 - \mathbf{x}_2)} \sqrt{(\mathbf{x}_1 - \mathbf{x}_2, \mathbf{x}_1 - \mathbf{x}_2)} = \|\mathbf{x}_1 - \mathbf{x}_2\| \|\mathbf{y}\| \end{aligned}$$

The above shows that the map $\mathbf{x} \rightarrow (\mathbf{x}, \mathbf{y})$ is uniformly continuous and therefore continuous. Since R is a continuous function and the composition of two functions is continuous we have that $L_h \mathbf{f}$ is continuous. \square

Most transforms are useless if one cannot invert them. Therefore, in order for all this machinery to have a purpose, we need to come up with an inversion formula for $L_h \mathbf{f}$. We will proceed by showing that L_h is an isometry from \mathcal{H} to $\mathcal{L}^2(\mathbf{G}, d\mu)$. Recall that by previous remarks, an isometry is invertible on its image. Thus, our second step will be to derive the appropriate inversion formulas. In order to show that L_h is an isometry, we need to use the following technical result first stated in [10].

Theorem 1.2. Let R be a square-integrable representation acting on \mathcal{H} . Then there exists in $U(\mathcal{H})$ a unique self-adjoint operator C such that the following hold:

1. The set of admissible vectors coincides with the domain of C .
2. Let \mathbf{h}_1 and \mathbf{h}_2 be any two admissible vectors. Let \mathbf{f}_1 and \mathbf{f}_2 be any two vectors in \mathcal{H} . Then:

$$\int_{g \in G} \overline{(R(g)\mathbf{h}_1, \mathbf{f}_1)} (R(g)\mathbf{h}_2, \mathbf{f}_2) d\mu = (C\mathbf{h}_1, C\mathbf{h}_2) (\mathbf{f}_1, \mathbf{f}_2)$$

3. If the group G is unimodular, then C is a multiple of the identity.

The proof of this theorem is quite technical and can be found in [10]. However, we will point out the assumptions that this proof uses and give a very rough sketch of some of the ideas involved. The proof is of a constructive nature – its general aim is to construct an operator C which satisfies the properties described above. The first step is to define an operator T_h for any admissible vector $\mathbf{h} \in \mathcal{H}$. Its domain is the set of vectors $\mathbf{f} \in \mathcal{H}$ such that:

$$\int_{g \in G} (R(g)\mathbf{h}, \mathbf{f}) d\mu < \infty$$

It is necessary to show that T_g is closed and bounded. For this purpose, one needs to use the fact shown in theorem 4.3, namely that the span of the set $\{R(g)\mathbf{h}\}$ is dense in \mathcal{H} since R is irreducible. Armed with this operator, the second step is to define another operator $C_{h_1 h_2}$ by $T_{h_1}^* T_{h_2}$. It can be shown that $C_{h_1 h_2}$ satisfies:

$$\int_{g \in G} \overline{(R(g)\mathbf{h}_1, \mathbf{f}_1)} (R(g)\mathbf{h}_2, \mathbf{f}_2) d\mu = C_{h_1 h_2}(\mathbf{f}_1, \mathbf{f}_2) \quad (3)$$

Finally using some considerations from the theory of linear operators one can show that there exists a unique positive operator C such that $C_{h_1 h_2} = (C\mathbf{h}_1, C\mathbf{h}_2)$ and that has all the desired properties. Combining this with 3 gives the desired result. We now proceed to state and prove the main theorem.

Theorem 1.3. Let $\mathbf{f} \in \mathcal{H}$. Then the map $\mathbf{f} \rightarrow L_h \mathbf{f}$ is an isometry from \mathcal{H} into $\mathcal{L}^2(G, d\mu)$.

Proof. We will invoke theorem 1.2 immediately. Consider the special case when $\mathbf{f}_1 = \mathbf{f}_2 = \mathbf{h}_1 = \mathbf{h}_2$. Then by part 2 of theorem we have that for any admissible vector h :

$$(C\mathbf{h}, C\mathbf{h}) = \frac{1}{\|\mathbf{h}\|^2} \int_{g \in G} |(R(g)\mathbf{h}, \mathbf{h})|^2 d\mu$$

Using this and letting $\mathbf{h} = \mathbf{h}_1 = \mathbf{h}_2$, by another application of theorem 1.2 we get that:

$$\int_{g \in G} \overline{(R(g)\mathbf{h}, \mathbf{f}_1)} (R(g)\mathbf{h}, \mathbf{f}_2) d\mu = (C\mathbf{h}, C\mathbf{h})(\mathbf{f}_1, \mathbf{f}_2) = \frac{\int_{g \in G} |(R(g)\mathbf{h}, \mathbf{h})|^2 d\mu}{\|\mathbf{h}\|^2} (\mathbf{f}_1, \mathbf{f}_2)$$

Now the desired result easily follows, since:

$$\begin{aligned} (\mathbf{f}_1, \mathbf{f}_2) &= \frac{\int_{g \in G} |(R(g)\mathbf{h}, \mathbf{h})|^2 d\mu}{\|\mathbf{h}\|^2} \int_{g \in G} \overline{(R(g)\mathbf{h}, \mathbf{f}_1)} (R(g)\mathbf{h}, \mathbf{f}_2) d\mu \\ &= \frac{1}{c_h} \int_{g \in G} \overline{(R(g)\mathbf{h}, \mathbf{f}_1)} (R(g)\mathbf{h}, \mathbf{f}_2) d\mu = \int_{g \in G} \overline{(L_h \mathbf{f}_1)(g)} (L_h \mathbf{f}_2)(g) d\mu \\ &= (L_h \mathbf{f}_1, L_h \mathbf{f}_2) \end{aligned}$$

□

Since L_h is an isometry, it is unitary map from $\mathcal{H} \rightarrow L_h(\mathcal{H}) \subset \mathcal{L}^2(\mathbf{G}, d\mu)$ which is its range and can be inverted there. In particular, let $\Psi \in L_h(\mathcal{H})$ then the inverse of L_h on its range is given by:

$$[L_h^{-1}\Psi](g) = \frac{1}{\sqrt{c_h}} \int_{g \in G} \Psi(g)(R(g)\mathbf{h})d\mu$$

We provide a sketch of the proof. In particular, we make no comment about the convergence of certain integrals that we present. For a complete proof, see [10].

Proof. The first thing we need to do is to find a precise description of which functions fall in the range of L_h and which ones do not. Armed with this description, the inversion formula will pop out immediately. Consider the following characterization of the range. Let $\mathbf{h} \in \mathcal{H}$ be an admissible vector for a square-integrable representation R . Define p_h :

$$p_h(g) = \frac{1}{c_h} (R(g)\mathbf{h}, \mathbf{h})$$

Then, it can be shown that for any $\Psi \in \mathcal{L}^2(\mathbf{G}, d\mu)$, Ψ is in the range of L_h iff for all $g \in G$ the following equation holds:

$$\Psi(g) = \int_{a \in G} p_h(a^{-1}g) \Psi(a) d\mu$$

At this point, the formula for the inverse is not too hard to derive. For if Ψ satisfies the above condition, it must be invertible since it is in the range of L_h , which is an isometry. Let $y \in G$.

$$\begin{aligned} L_h \left(\frac{1}{\sqrt{c_h}} \int_{g \in G} R(g)\mathbf{h} \Psi(g) d\mu \right) (y) &= \frac{1}{\sqrt{c_h} \sqrt{c_h}} \int_{g \in G} \Psi(g) (R(y)\mathbf{h}, R(g)\mathbf{h}) d\mu \\ &= \frac{1}{c_h} \int_{g \in G} \Psi(g) (R(g^{-1}y)\mathbf{h}, \mathbf{h}) d\mu = \int_{g \in G} \Psi(g) p_h(g^{-1}y) d\mu = \Psi(y) \end{aligned}$$

Notice that we have used the fact that R is unitary. Therefore, multiplying both sides by L_h^{-1} (we can do this since Ψ is in the range of L_h) we get the precise inversion formula:

$$[L_h^{-1}\Psi](g) = \frac{1}{\sqrt{c_h}} \int_{g \in G} \Psi(g) R(g)\mathbf{h} d\mu$$

□

Here, we conclude our discussion of the properties of L_h . There are many more that we could have talked about, but the primary ones were touched upon. Most importantly, we have presented a general framework within which to construct integral transforms and analyze them. Whenever faced with an abstraction it is important to seek out examples. Therefore, it is highly recommended that the reader skip ahead to the next chapter which has a number of explicit constructions within the described framework worked out. The theory in the remainder of this chapter will not be used until chapter 7, and so it will be “safe” to return to it after chapter 6.

2 Extension to Homogeneous Spaces

After reading the previous section, a natural question to ask is whether or not one could execute the construction presented for any locally compact group G and Hilbert Space \mathcal{H} . Indeed, the only thing we know for certain is that we can always find a left Haar measure on G . However, we are not always guaranteed to find a square-integrable representation R of G in $GL(\mathcal{H})$. Unfortunately, there are some instances when it would be really useful to be able to define the transform anyway, perhaps by sacrificing some restriction, but still getting most of the implications. We describe such a generalization in this section. Similarly to the last section, we hold off the examples until Chapter 7 where we explain the construction of a Continuous Wavelet Transform on the 2-Sphere. This discussion closely follows that of [2].

Recall that the usual three steps in the construction are to find a left Haar measure on a given locally compact group G , to find a continuous unitary cyclic (or irreducible) representation R and, finally, to show that it is square-integrable by finding an admissible vector $h \in \mathcal{H}$. The point of failure is the fairly stringent requirement of square-integrability. For any given G and \mathcal{H} , one may not always be able to find $R : G \rightarrow GL(\mathcal{H})$ such that R is square-integrable. However, it may be possible to show that a restriction of R to a homogeneous space is square-integrable. In what follows, we discuss the details of this restriction and its implications for the general framework from the previous section. We begin by defining what is meant by a homogeneous space.

Definition 2.1. Let G be a group and X a set. We say that G acts on X *transitively* from the left if for every $x, y \in X$ there exists $g \in G$ such that $gx = y$. If it is possible to define a left transitive action of G on X , we say that X is a *transitive G -space*. As stated in [8].

Definition 2.2. Let G be a locally compact group and X be a transitive G -space. We say that X is a *homogeneous space* if there is a closed subgroup H of G such that X is homeomorphic to G/H .

A few comments about the definition above. When we say that a subgroup of a locally compact group is closed we mean that the underlying topological space of H is a closed subspace of G and not that H is closed in the group-theoretical sense. Also a simple application of theorem 2.2 shows that X is locally compact and therefore there exists a left Haar measure on X .

Example 2.1. Consider the locally compact group $SO(3)$. Clearly $SO(2)$ is a closed subgroup and therefore the 2-Sphere is a homogeneous space since \mathbb{S}^2 is homeomorphic to $SO(3)/SO(2)$. A left (and right) Haar measure on \mathbb{S}^2 is given by $\sin(\theta)d\theta d\phi$ where \mathbb{S}^2 is parameterized by standard spherical coordinates as described in chapter 2.

Let us now proceed to explain how to restrict a representation R from the group G to a homogeneous space $X = G/H$. We will then discuss what is lost by this restriction. We need one more definition to proceed.

Definition 2.3. Let $X = G/H$ and $e, g \in G$ and where e denotes the identity element. For each coset $gH \in X$ define the function $\phi : G/H \rightarrow G$ that acts by mapping the coset gH to any element in it. Then we define the function $\sigma : G/H \rightarrow G$ by:

$$\sigma(gH) = (\phi(eH))^{-1}\phi(gH)$$

When ϕ is chosen in such a way that σ is Borel measurable (see Theorem 2.2) then the function σ is called a *Borel section*. It is not uniquely defined. (Adapted from [3]).

The above definition is quite technical, but the most important idea behind it is that the Borel section is a map lifting elements from the homogeneous space X back up to the group G . By using this map we can define a representation in a slightly more generalized way. The representation will not be a function of all elements in the group G , but only on those elements in G that are in the image of a Borel section of some homogeneous space X .

To make this formal, let $\sigma : G/H = X \rightarrow G$ be a Borel section and ν a left Haar measure on X . Then the generalized definition for square-integrability and admissibility is the following. Let $\mathbf{h} \in \mathcal{H}$ and R an irreducible representation of G in $U(\mathcal{H})$. We say that \mathbf{h} is admissible and R is square-integrable if the following holds for all $\mathbf{f} \in \mathcal{H}$:

$$\int_{x \in X} |(R(\sigma(x))\mathbf{h}, \mathbf{f})|^2 d\nu < \infty \quad (4)$$

Notice that we recover the original definition of square-integrability if we let H be the trivial group. Putting this fancy generalization aside one may wonder how this relaxation of the definition of square-integrability affects the general framework described in the previous section. The immediate answer is that the resulting transform will not be a function of all the elements of a group G , but only those in the image of the Borel section σ . This may very well limit the analyzing power of the resulting transform. Nevertheless, it does not have to be a setback in all cases. It may be possible to start with a group larger than what is necessary and design the homogeneous space X and Borel section σ so that precisely the relevant group elements are in its image. In fact, this is precisely what is done in the derivation of the Continuous Wavelet Transform on the 2-Sphere presented in Chapter 7.

Examples of Integral Transforms

In this chapter we present two important examples of the left integral transform that was introduced in the last chapter. In particular, we will derive the Windowed Fourier Transform (WFT) and the Continuous Wavelet Transform (CWT). It is exciting that both of these transforms, as well as others, arise in the same context for two reasons. The first is that both transforms are extensively used in signal processing, computer vision, graphics, physics and a myriad of other applications. Therefore, being able to understand their group theoretical background is incredibly useful. The second reason for finding the construction exciting is that these two transforms have completely different properties and uses. In a sense, it is unexpected that two transforms of a nature so different should arise in the same context.

In deriving these transforms, we will follow the framework in the last chapter very closely and sequentially. This means that in order to present an example, we will first pick a locally compact group G . Secondly, we will give a left Haar measure for this group and verify, in some cases, that it is indeed invariant under the action of G . The third step will be to pick a Hilbert Space \mathcal{H} of interest and to find a square-integrable representation of G in $GL(\mathcal{H})$. Of course, we will have to explicitly pick the analyzing vector and justify the choice, since most of the time there are many to choose from. The actual transforms then will be written down as well as their inverses. We will also present the derivation for the 2-dimensional CWT as the construction varies somewhat from the one-dimensional case.

In reading the examples presented, the reader should try to focus more on the derivation rather than on the properties of the transforms themselves. There is a great deal of literature available about the Continuous Wavelet Transform and the Windowed Fourier Transform, but much less is available about methods used to construct transforms for ad hoc purposes. The group-theoretical framework presented in the last chapter is general enough to allow a large number of specific transforms to be defined and it is important that one is able make use of this. Let us do so now for the WFT and CWT.

1 Continuous Wavelet Transform on \mathbb{R}

We begin with the CWT because the underlying group is somewhat simpler than the one used for the derivation of the WFT.

We begin the construction of the Continuous Wavelet Transform by letting G be the group of affine transformations of the real line \mathbb{R} . Let us

discuss the group in some detail. Some of the material presented can be found in [1] and [14].

Definition 1.1. Let $a \in R_*^+$ and $b \in \mathbb{R}$. Then an *affine transformation* of the real line $T_{a,b} : \mathbb{R} \rightarrow \mathbb{R}$ is defined by :

$$T_{a,b}(x) = ax + b$$

For this reason G is commonly called the $ax + b$ group. Since each transformation is determined by a and b , we write each element in G as (a, b) . The group law \circ is defined by $(a, b) \circ (a', b') = (aa', b + ab')$. This is quite natural since:

$$a(a'x + b') + b = aa'x + ab' + b = (aa')x + (ab' + b)$$

The identity element is $(1, 0)$ and the inverse of (a, b) is given by $(\frac{1}{a}, -\frac{b}{a})$. The $ax + b$ group can also be identified with the group of matrices $\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$ with matrix multiplication as the group law. Finally, the $ax + b$ group can be identified with $R \rtimes R_*^+$, where \rtimes is the semidirect product defined in equation(19).

The $ax + b$ group is not unimodular because its left and right Haar measures are not the same. It turns out that the left Haar measure is $\frac{dad b}{a^2}$, while the right Haar measure is $\frac{dad b}{a}$. Let τ denote a left action of the group. By definition:

$$\tau(a, b) : (a', b') \rightarrow (aa', ab' + b)$$

The Jacobian of τ is $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$. Then by an application of theorem 4.2 we have that the left Haar measure on the $ax + b$ group is given by $\frac{dad b}{a^2}$. In applying this theorem be careful not to confuse the affine maps defined in the statement of the theorem with the affine maps that are part of the $ax + b$ group because they are different. In the statement of the theorem, $A(x)$ is equivalent to the Jacobian of τ .

In signal processing, the continuous wavelet transform is a tool that is used to analyze finite energy signals. Thus, we will employ $\mathcal{L}^2(\mathbb{R}, dx)$ as the underlying Hilbert Space. Next, we need to find a suitable unitary representation R of the $ax + b$ group on $U(\mathcal{L}^2(\mathbb{R}, dx))$. Keep in mind that a representation is a homomorphism and accordingly not all of $U(\mathcal{L}^2(\mathbb{R}, dx))$ has to be in the image of R — only a closed subspace. For any $f \in \mathcal{L}^2(\mathbb{R}, dx)$, the natural choice for $R : G \rightarrow U(\mathcal{L}^2(\mathbb{R}, dx))$ is the following:

$$[R(a, b)f](x) = \frac{1}{\sqrt{a}}f\left(\frac{x - b}{a}\right) \quad (5)$$

It is easy to see that R defined in this way is continuous as a function on G and is unitary. It is also required that R be irreducible and the trouble is that the one above is not. However, recall that we need irreducibility only to show that the span of $R(a, b)\mathbf{f}$ is dense in $\mathcal{L}^2(\mathbb{R}, dx)$. It turns out that R is a cyclic representation [1] and thereby satisfies this property without being irreducible. Thus, this is the example of a cyclic non-irreducible representation that was promised before.

Now all that remains is to show that R is square-integrable by choosing an admissible analyzing wavelet $\psi \in \mathcal{L}^2(\mathbb{R}, dx)$. Recall that a wavelet ψ is admissible if

$$M = \int_{(a,b) \in G} |(R(a, b)\psi, \psi)|^2 \frac{dad b}{a^2} < +\infty \quad (6)$$

When $\psi \in \mathcal{L}^1(\mathbb{R}, dx) \cap \mathcal{L}^2(\mathbb{R}, dx)$ then it can be shown that the admissibility condition is equivalent to the requirement that $\hat{\psi}(0) = 0$. As we know from the previous section, functions ψ that have this property are known as wavelets. We are getting quite close now. Indeed, we can now define the left integral transform for any function $\mathbf{f} \in \mathcal{L}^2(\mathbb{R}, dx)$. We will denote it by W_ψ because it is in fact the Continuous Wavelet Transform. Assign to ψ the positive number c_ψ defined by:

$$c_\psi = \frac{1}{\|\psi\|^2} \int_{(a,b) \in G} |(R(a, b)\psi, \psi)|^2 \frac{dad b}{a^2} \quad (7)$$

Then the Continuous Wavelet Transform is given by:

$$(W_\psi f)(a, b) = (R(a, b)\psi, \mathbf{f}) = 1/\sqrt{ac_\psi} \int_{-\infty}^{+\infty} \psi\left(\frac{x-b}{a}\right) f(x) dx \quad (8)$$

We proved in the last chapter that W_ψ is an isometry and is therefore invertible. A function $\mathbf{f} \in \mathcal{L}^2(\mathbb{R}, dx)$ can be reconstructed by the formula:

$$f(x) = 1/\sqrt{c_\psi} \iint_{R_+^2} \psi\left(\frac{x-b}{a}\right) ((W_\psi f)(a, b)) \frac{dad b}{a^2} \quad (9)$$

We have successfully derived exactly the Continuous Wavelet Transform as the one that was presented in the previous section.

2 Windowed Fourier Transform

The group underlying the Windowed Fourier Transform is the Weyl-Heisenberg group. Let us spend a few moments and describe its properties. The interested reader should refer to [11] and [1] for more details. Let $s, b \in \mathbb{R}$, and $\tau \in \mathbb{C}$ so that $|\tau| = 1$. An element of the Weyl-Heisenberg group (G_{WH} for short) can be written as a triplet $\gamma = \{s, b, \tau\}$ with the following group law:

$$\gamma \circ \gamma' = \{s + s', b + b', \tau\tau' e^{i(s'b - b's)/2}\}$$

Inverses are given by a more readable expression:

$$\gamma^{-1} = \{-s, -b, \tau^{-1}\}$$

Since $|\tau| = 1$, we can write $\tau = e^{ix\phi}$. Then left and right Haar measure on G_{WH} are both given by $dsdbd\phi$. This is not too hard to verify by using theorem 4.2. Since the left and right Haar measures are the same, G_{WH} is a unimodular group. Let us try to give some intuition about how this group arises. There are two places where the Weyl-Heisenberg group comes up: quantum mechanics and signal processing. Even though, the quantum mechanical setting is more natural — in the context of this paper, let us attempt to understand the group from the signal processing viewpoint. In particular, the group appears when one considers shifts in the time and frequency domains simultaneously for some finite energy signal $\mathbf{f} \in \mathcal{L}^2(\mathbb{R})$. However, from Fourier analysis we know that shifting a signal in the frequency domain is equivalent to modulating the signal in the time domain. Thus, in analyzing \mathbf{f} we will need to define two operators – one for translation and one for modulation.

Definition 2.1. Let $s \in \mathbb{R}$ and $\mathbf{f} \in \mathcal{L}^2(\mathbb{R}, dx)$. Let T^s denote the *translation operator* defined by:

$$(T^s \mathbf{f})(x) = \mathbf{f}(x - s)$$

Definition 2.2. Let $b \in \mathbb{R}$ and $\mathbf{f} \in \mathcal{L}^2(\mathbb{R}, dx)$. Let E^b denote the *modulation operator* defined by:

$$(E^b \mathbf{f})(x) = e^{ibx} \mathbf{f}(x)$$

The continuous wavelet transform allows us to study a signal by varying its scale and shifting it. Similarly, using the translation and modulation operators, we can study a signal by shifting and modulating it. To do this let us define a transformation $W(s, b)$ that combines translations and dilations by:

$$W(s, b) = e^{\frac{isb}{2}} T^s E^b \quad (10)$$

It can easily be shown that the operators $W(s, b)$ satisfy the following equality:

$$W(s, b)W(s', b') = e^{\frac{i(s'b - b's)}{2}} W(s + s', b + b') \quad (11)$$

It is certainly not a coincidence that this is exactly the law of composition of the Weyl-Heisenberg group. From equation (11) it immediately follows that the map $R : \gamma \rightarrow \tau W(s, b)$ is a group homomorphism. In fact we have more — by setting \mathcal{H} to $\mathcal{L}^2(\mathbb{R}, dx)$ in our general framework, we find that the map: $R : \gamma \rightarrow \tau W(s, b)$ is a continuous, irreducible, unitary representation of G_{WH} in $U(\mathcal{L}^2(\mathbb{R}, dx))$. Thus, equipped with a locally compact group G_{WH} with an appropriate left Haar measure, a Hilbert Space $\mathcal{L}^2(\mathbb{R}, dx)$ and an appropriate representation R of G_{WH} , we only need to verify that R is square-integrable before we can write down the Windowed Fourier Transform explicitly.

It turns out that because G_{WH} is unimodular, every vector in $\mathcal{L}^2(\mathbb{R}, dx)$ is admissible. For a proof of this see [23]. For any admissible vector ψ and all $\mathbf{f} \in \mathcal{L}^2(\mathbb{R}, dx)$ we can define the Windowed Fourier Transform WF_ψ . Assign to ψ the positive number c_ψ defined by::

$$c_\psi = \frac{1}{\|\psi\|^2} \int_{\gamma \in G} |(\tau W(s, b)\psi, \psi)|^2 dsdbd\theta \quad (12)$$

Then the Windowed Fourier Transform is given by:

$$WF_\psi(s, b) = 1/\sqrt{c_\psi} (\tau W(s, b)\psi, \mathbf{f}) = 1/\sqrt{c_\psi} \int_{-\infty}^{\infty} e^{i(b+\theta)x} \psi(x-s) f(x) dx \quad (13)$$

The inversion formula is given by:

$$f(x) = 1/\sqrt{c_\psi} \iint [W(s, b)\psi](x) WF_\psi(s, b) dsdb \quad (14)$$

Thus, we have presented the full group-theoretical derivation of the one dimensional Windowed Fourier Transform. Generalizing this construction to functions defined in $\mathcal{L}^2(\mathbb{R}^n)$ is not very difficult. In fact, one only has to increase the dimension of the variables s and b to the dimension of the domain of the signal and the above construction works in the same way.

The generalization for the Continuous Wavelet Transform is somewhat more tricky and we deal with it next.

3 Continuous Wavelet Transform on \mathbb{R}^2 and beyond

Since the detailed derivation of two transforms has already been presented, we proceed quickly in the derivation of the Continuous Wavelet Transform. The group underlying the derivation, in this case, is the similitude group of the plane, denoted by $SIM(2)$. It consists of three operations on the plane – translations, dilations and rotations. Let $a \in \mathbb{R}^+$, $\mathbf{b} \in \mathbb{R}^2$, $\theta \in [0, 2\pi)$, then an element $g \in SIM(2)$ is a triplet (a, \mathbf{b}, θ) . The similitude group of the plane is formally defined as a semidirect product:

$$SIM(2) = \mathbb{R}^2 \rtimes (R_*^+ \times SO(2))$$

The group law is implicit in the above construction. Left Haar measure on $SIM(2)$ is given by $\frac{dad^2b d\theta}{a^3}$. Since the 2D CWT is used to analyze functions of two variables, the Hilbert Space that we will use in the construction is $U(\mathcal{L}^2(\mathbb{R}^2, d\mathbf{x}))$.

Let r_θ denote the usual 2×2 rotation matrix in $SO(2)$. Then a continuous, unitary, irreducible representation $R : G \rightarrow \mathcal{L}^2(\mathbb{R}^2, d\mathbf{x})$ is given by:

$$[R(g)f](\mathbf{x}) = [R(a, \mathbf{b}, \theta)f](\mathbf{x}) = \frac{1}{a} f\left(\frac{r_{-\theta}(\mathbf{x} - \mathbf{b})}{a}\right)$$

The admissibility condition for an analyzing vector $\psi \in \mathcal{L}^2(\mathbb{R}^2, d\mathbf{x})$ becomes:

$$\int_{g \in SIM(2)} |(R(g)\psi, \psi)|^2 dg < \infty$$

Similar to the case of the 1D CWT, the admissibility condition reduces to $\hat{\psi}(0) = 0$ if $\psi \in \mathcal{L}^1(\mathbb{R}^2, d\mathbf{x}) \cap \mathcal{L}^2(\mathbb{R}^2, d\mathbf{x})$. We are now ready to define the 2D CWT. Assign to an admissible vector ψ the positive number c_ψ defined by:

$$c_\psi = \frac{1}{\|\psi\|} \int_{g \in SIM(2)} |(R(g)\psi, \psi)|^2 dg.$$

By taking Fourier Transforms, one gets an equivalent expression:

$$c_\psi = (2\pi)^2 \int |\hat{\psi}(k)|^2 \frac{d\mathbf{k}}{|k|^2}$$

Thus, given an admissible wavelet ψ , the 2D CWT is defined by:

$$W_\psi(a, \mathbf{b}, \theta) = \frac{1}{a} \int \overline{\psi\left(\frac{r_\theta(\mathbf{x} - \mathbf{b})}{a}\right)} f(\mathbf{x}) d\mathbf{x}$$

Its inverse is given by

$$f(\mathbf{x}) = \frac{1}{c_\psi} \iiint W_\psi(a, \mathbf{b}, \theta) \psi\left(\frac{r_\theta(\mathbf{x} - \mathbf{b})}{a}\right) \frac{da d\mathbf{b} d\theta}{a^3}$$

Signal Processing on the 2-Sphere

1 Introduction

We approach the subject matter in this chapter from a different angle. The main focus is the development of the Continuous Wavelet Transform on the 2-sphere. However, it is important to first provide sufficient motivation for the spherical Continuous Wavelet Transform as it requires the use of the sophisticated machinery of representations of homogeneous spaces. Indeed, it is true that the CWT on the sphere has an important role to play, one that has not yet been filled by other methods. Moreover, it should be noted that in the last decade, the computer graphics and computer vision communities have posed challenges and questions, to resolve, which it seems necessary to develop signal processing tools on the 2-sphere. Due to these considerations, we present other techniques in Spherical Signal Processing in Appendix C.

It should be mentioned that we will not discuss any discrete signal processing techniques on the sphere and focus solely on continuous methods. There are two separate reasons for this decision. First, in order to discuss discrete methods we would need to introduce a large amount of background material such as the theory of frames and sampling theory, which would make this paper unreasonably long. Second, not all discrete transforms arise from the discretization of continuous ones. In fact, their derivation is completely unrelated to the theory of continuous transforms that we have presented thus far and therefore is best discussed in a separate paper. Nonetheless, continuous methods play a very important and practical role in many applications and their discussion will be our main focus in the following sections. Accordingly, in this chapter we derive a local integral transform using the previously developed group-theoretical framework and motivate this derivation with signal processing considerations.

To begin, let us describe two classes of techniques, global and local, that include all of the techniques we discuss. The names of the classes speak for themselves. Global techniques are best used for studying properties that appear throughout the entire signal. On the other hand, local techniques are useful for studying properties that only occur on part of the signal. All of the transforms presented in the last chapter are examples of local techniques, and each of them is designed to focus on part of a signal in a particular way. We leave the discussion of Global Techniques until Appendix C and begin with a discussion of a local transform - the Spherical Continuous Wavelet Transform.

2 Wavelet Transform on the 2-Sphere

The derivation in this section is due to Antoine and Vandergheynst [2]. In designing a transform that is to be used for analyzing a signal locally, we are faced with an important decision about what properties we want our transforms to have. Let us recall two of the transforms presented in the last chapter. The windowed Fourier transform analyzes the frequency content of signals by varying the frequency of the analyzing function within a window of fixed width. It is efficient at detecting frequencies, in particular ones that are low. On the other hand, the continuous wavelet transform analyzes the frequency content of signals by varying the scale of the analyzing function of constant frequency. It is an excellent singularity detector. While there are constructions for both types of transforms on the sphere, in this paper, we present the derivation of the spherical wavelet transform because it has more immediate applications to problems in omnidirectional vision than the windowed Fourier approach.

The construction is fairly sophisticated, so we break the discussion up into two parts. In the first, we discuss some preliminary material dealing with transformations on the sphere and give some motivation for how we want to construct the transform. In the second section, we present highlights of the actual construction.

2.1 Preliminaries

First, we need to clarify which transformations we are interested in using to analyze signals on the sphere. Since we would like to develop a wavelet-like transform we would need to describe three transformations that are traditionally used in wavelet analysis: translations, rotations and dilations.

Translations are easy to identify for they correspond to rotations about the θ and ϕ axes. Rotations can be realized as rotations about a certain axis on the 2-sphere as well. Thus, translations and rotations together can be associated with the action of $SO(3)$ on \mathbb{S}^2 . Then for each $g \in SO(3)$ recall the unitary rotation operator:

$$[\Lambda_g f](\eta) = f(g^{-1}\eta)$$

Dilation or scaling is more difficult to define. Since \mathbb{S}^2 is compact, dilating directly on it leads to problems. Let $a \in \mathbb{R}_*^+$ be the scaling parameter. The most naive approach would be to define the dilation operator D_a on

$\mathcal{L}^2(\mathbb{S}^2, d\mu)$ in the following fashion:

$$[D_a f](\eta) = f(\eta_a)$$

where $\eta_a = (\cos(\phi) \sin(\theta_a), \sin(\phi) \sin(\theta_a), \cos(\theta_a))$ and $\theta_a = a\theta$. However, a problem occurs when $a > 1$. In this case, some functions will be stretched beyond the circumference of the sphere, which will lead to the function overlapping itself. This means that at some points on the domain, the function will take on multiple values. Therefore, dilating in this manner is ill-defined and we are forced to use a slightly more complicated method.

An idea proposed in [2] is to dilate functions on the plane tangent to sphere at its North Pole with the aid of the stereographic projection. To dilate a function about some point η_0 on the sphere, three steps are needed. First η_0 is rotated to the North pole by a rotation R . Next, the function is stereographically projected to the plane tangent to the North pole and dilated there in the same way that dilations are performed for the regular 2D CWT. Finally, the result is lifted back up to the sphere using the inverse stereographic projection and rotated back by R^{-1} . Defining dilations in this way avoids the previously mentioned difficulties — and is well-defined. It can be shown through geometrical considerations that the dilation operator thus defined acts in the following way:

$$[D_a f](\eta) = f(\eta_a)$$

where $\eta_a = (\cos(\phi) \sin(\theta_a), \sin(\phi) \sin(\theta_a), \cos(\theta_a))$ and $\tan(\frac{\theta_a}{2}) = a \tan(\frac{\theta}{2})$.

Having described the transformations necessary for constructing our transform, we need to begin to apply the group-theoretical framework developed in chapter 5. As always, the first step is to find a group whose square-integrable representation in $U(\mathcal{L}^2(\mathbb{S}^2, d\mu))$ will give rise to the transform we want. However, the trouble is that there is no easy way to find a group whose elements correspond in some way to both $SO(3)$ and \mathbb{R}_*^+ . The direct product $SO(3) \times \mathbb{R}_*^+$ is too large and has no useful relationship to the transformations we are concerned with. It can also be shown that no semi-direct product between the two groups is possible. So we are left with no way to make $SO(3)$ and \mathbb{R}_*^+ into a group and therefore no way to continue our construction. Therefore, we must fall back onto the more sophisticated machinery of representations of homogeneous spaces. In particular, if it were possible that $SO(3)$ together with \mathbb{R}_*^+ could be considered a part of a homogeneous space for some locally compact G and a subgroup H then we would have a way out. This was the main observation made in [2].

2.2 The Construction

First, let us set down some notation. Let G and H be groups and let $G \cdot H$ denote the set of products gh where $g \in G$ and $h \in H$. The multiplication between elements of the groups G and H is understood to be well-defined if the notation is to be used.

The solution proposed in [2] was to consider the space $SO(3) \cdot \mathbb{R}_*^+$ as part of the Lorentz group $SO(3, 1)$, where elements in $SO(3)$ and \mathbb{R}_*^+ are written as 4×4 matrices. Let us give a brief outline for the construction with this in mind. First, we will show that there exists a subgroup H of $SO(3, 1)$ such that $SO(3, 1)/H = SO(3) \cdot \mathbb{R}_*^+$ is a homogeneous space. Next we will find an irreducible representation of $SO(3, 1)$ in $U(\mathcal{L}^2(\mathbb{S}^2, d\mu))$ such that it is square-integrable when restricted to $SO(3, 1)/H$. Using this, we state an admissibility condition for vectors in $\mathcal{L}^2(\mathbb{S}^2, d\mu)$. Lastly, according to the general framework, we write down the Continuous Wavelet Transform on the 2-Sphere alongside with its inverse.

To begin, we briefly discuss $SO(3, 1)$, the Lorentz group. Let $I_{3,1}$ denote the following matrix:

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}$$

Definition 2.1. The set of transformations $\{A \in GL_4(\mathbb{R}) \mid A^t I_{3,1} A\}$ is the stabilizer of $I_{3,1}$ is called the Lorentz group and is denoted by $SO(3, 1)$. (As stated in [18])

The following property of the Lorentz group is crucial to us. It can be shown that $SO(3, 1)$ can be written as a product of three groups as follows:

$$SO(3, 1) = SO(3) \cdot \mathbb{R}_*^+ \cdot \mathbb{C} \quad (15)$$

This decomposition is called the *Iwasawa Decomposition* and it implies that each element $g \in SO(3, 1)$ can be written uniquely as a product of three elements α, β, γ where $\alpha \in SO(3)$, $\beta \in \mathbb{R}_*^+$ and $\gamma \in \mathbb{C}$. This multiplication is well-defined when we write elements of each group as 4×4 matrices. One can obtain the exact matrix form of the above product via the derivation of the Iwasawa Decomposition, however, this is beyond the scope of this paper and will not be presented here. What is important to realize is that we now have a way to look at $SO(3)$ and \mathbb{R}_*^+ as part of the Lorentz group.

To show that $SO(3) \cdot \mathbb{R}_*^+$ is a homogeneous space, consider all of the elements $g \in SO(3, 1)$ that can be written as: $\alpha \cdot \beta \cdot \gamma$ where α and β are the identity elements in $SO(3)$ and \mathbb{R}_*^+ , respectively. Clearly this set is

isomorphic to \mathbb{C} and is a subgroup of $SO(3, 1)$. Then, by taking quotients in equation (15) we have the following:

$$X = SO(3, 1)/\mathbb{C} = SO(3) \cdot \mathbb{R}_*^+$$

The above formula states that all elements in the quotient space $SO(3, 1)/\mathbb{C}$ can be written as products of two elements $\alpha \in SO(3)$ and $\beta \in \mathbb{R}_*^+$. It can be shown that $SO(3, 1)$ acts transitively on X and as therefore X is a homogeneous space, by definition. We can now proceed to define a square-integrable representation of X on $U(\mathcal{L}^2(\mathbb{S}^2, d\mu))$ and derive a useful transform as planned.

Let us specify the usual elements in our construction of a representation of a group. The Hilbert Space we are interested in is, of course, $\mathcal{L}^2(\mathbb{S}^2, d\mu)$. We write elements of X as pairs (α, β) where $\alpha \in SO(3)$ and $\beta \in \mathbb{R}_*^+$. Finally, let $d\nu$ denote the $SO(3, 1)$ left-invariant measure on X .

Then the next step according to the general theory is to find a continuous irreducible representation $R : SO(3, 1) \rightarrow U(\mathcal{L}^2(\mathbb{S}^2, d\mu))$. One such representation is given by:

$$[R(g)f](\eta) = \lambda(g, \eta)f(g^{-1}\eta), \quad \text{where } g \in SO(3, 1). \quad (16)$$

The factor $\lambda(g, \eta)$ is inserted because $d\mu$ is invariant under actions of only $SO(3)$ and not all of $SO(3, 1)$. However, the measure on $\mathcal{L}^2(\mathbb{S}^2)$, multiplied by this factor is invariant under left action by $SO(3, 1)$. Here, we give its explicit form without the derivation.

$$\lambda(g, \eta) = \frac{4a^2}{[(a^2 - 1)\cos(\theta) + (a^2 + 1)]^2}$$

Next, we need to restrict this representation to the homogeneous space X . The Borel section $\sigma : X \rightarrow SO(3, 1)$ is quite naturally given by:

$$\sigma(\alpha, \beta) = \alpha \cdot \beta$$

Therefore, the representation (16) has the following form when restricted to the homogeneous space X :

$$[R(\sigma(\alpha, \beta))f](\eta) = \lambda(\sigma(\alpha, \beta), \eta)f(\sigma(\alpha, \beta)^{-1}\eta), \quad \text{where } \sigma(\alpha, \beta) \in SO(3, 1). \quad (17)$$

The admissibility condition for some vector $\psi \in \mathcal{L}^2(\mathbb{S}^2)$ is nothing more than the integral that is defined in Chapter 5 with the representation understood to be the one define above. We say that $\psi \in \mathcal{L}^2(\mathbb{S}^2)$ is admissible if:

$$\int_X |(R(\sigma(\alpha, \beta))\psi, \psi)|^2 d\nu \quad (18)$$

It turns out that there is a large number of admissible vectors and that in fact they are dense in $\mathcal{L}^2(\mathbb{S}^2), d\mu$. The final step is to define a continuous wavelet transform of a function $f \in \mathcal{L}^2(\mathbb{S}^2, d\mu)$ given an admissible vector ψ .

$$[W_\psi f](\alpha, \beta) = \int_{\mathbb{S}^2} (\overline{R(\sigma(\alpha, \beta))\psi}, f(\eta)) d\mu$$

According to the developed theory, this transform is invertible and its inverse is given by:

$$f(\eta) = \int_{R_*^+} \int_{SO(3)} [W_\psi f](\alpha, \beta) R(\sigma(\alpha, \beta))\psi d\nu$$

Conclusions

Integral transforms such as the Continuous Wavelet Transform and the Windowed Fourier Transform play an important role in signal processing. Even though they have completely different properties, it is possible to study them within the same framework, that of square-integrable group representations. Introduced in 1985 by Grossman et. al., this framework is an incredibly powerful tool using which one can construct many different types of integral transforms to study many aspects of signals defined on a variety of domains.

The method itself is not very complicated to use. One begins with a Hilbert Space \mathcal{H} of signals that are of interest and some group. The next step is to find a square-integrable representation of this group in the space of unitary operators acting on \mathcal{H} . Typically, the group employed is chosen in such a way that the image of the representations is a set of transformations which can be used to study the signals at hand in some useful way. Finally, one can choose an appropriate admissible vector and construct an integral transform which can then be directly applied to studying signals in a wide variety of ways.

The strength of this construction is that it does not depend on any particular Hilbert Space and/or group. As a result, one has the freedom to use the construction in many cases. We have already used a number of Hilbert Spaces such as $\mathcal{L}^2(\mathbb{R})$, $\mathcal{L}^2(\mathbb{R}^n)$ and $\mathcal{L}^2(\mathbb{S}^2)$. Clearly many more are possible. The group can be varied as well. In the case of the WFT and CWT on the $\mathcal{L}^2(\mathbb{R})$, the Hilbert Space was the same but the choice of a different group led to radically different transforms. Moreover, there is much freedom in the choice of an admissible vector. When the group underlying the construction is unimodular, as in the case of the Weyl-Heisenberg group, all vectors are admissible. Thus, one can design very different transforms depending on the choice of the admissible vector.

Hopefully reading this paper will enable the reader to construct integral transforms for their own purposes. While this might seem daunting, the method is relatively straightforward given the abstract framework that has been presented. The choices of the group G and Hilbert Space \mathcal{H} are frequently dictated by the application at hand. The true freedom comes from the choice of representation (if there is more than one to select from) and analyzing vector. While not all transforms can be created in this manner, the framework is indeed quite general and can be extended if necessary as we have seen in the case of the spherical wavelet transform.

Appendix A: Elementary Group Theory

In this appendix we present a brief review of group theory since it is heavily used throughout the text. For a more detailed reference, the reader may consult [18] or [16].

Definition 0.1. A *group* is an ordered pair $(G, *)$ where G is a set and $*$ is a binary operation on G satisfying the following four conditions:

1. Closure: If $a, b \in G$ then $a * b \in G$.
2. Associativity: If $a, b, c \in G$ then $a * (b * c) = (a * b) * c$
3. Identity Element: $\exists e \in G$ such that, $\forall a \in G, a * e = e * a = a$.
4. Inverses: $\forall a \in G, \exists a^{-1} \in G$ such that $a * a^{-1} = a^{-1} * a = e$.

In many instances when the group law is clear, we write G instead of the more cumbersome $(G, *)$.

Example 0.1. An important example of a group is the symmetric group on n elements denoted by S_n . It is the group of all permutations of the set $\{1 \dots n\}$. It is easy to see that the cardinality of $S_n = n!$.

Definition 0.2. Let $H \subseteq G$. We call H a *subgroup* if it is closed under the group law, contains the identity and contains an inverse for each element in it.

There are many types of subgroups one can consider in group theory. Throughout the appendix we will discuss a number of them including the one we define now: normal subgroups.

Definition 0.3. Let G be a group and H a subgroup. The subgroup H is said to be *normal* if for all $g \in G$ and $h \in H$, the element ghg^{-1} (called the conjugate) belongs to H . Alternatively, a subgroup H is *normal* if it is closed under conjugation by elements in G .

One way to study the structure of a group is to map a simpler group into it. Since we may know more about the structure of the simpler group, we will have found out something about the larger group. There are many maps of this type, let us discuss one here.

Definition 0.4. Let G and G' be groups. A map $\phi : G \rightarrow G'$ is called a *homomorphism* if for all $a, b \in G$ we have that $\phi(a)\phi(b) = \phi(ab)$.

Example 0.2. The most simple example of a homomorphism is as follows. Let G and G' be groups. Define the map $\phi : G \rightarrow G'$ by $\phi(g) = e'$ for all $g \in G$. We say that ϕ is the *trivial homomorphism*.

There are two other important subgroups that naturally arise alongside group homomorphisms.

Definition 0.5. Let G and G' be groups and $\phi : G \rightarrow G'$ a group homomorphism. The following set is called the *kernel* of the homomorphism ϕ :

$$\ker(\phi) = \{g \in G \mid \phi(g) = e\}$$

It is easy to verify that $\ker(\phi)$ is a subgroup of G .

Definition 0.6. Let G and G' be groups and $\phi : G \rightarrow G'$ a group homomorphism. The following set is called the *image* of the homomorphism ϕ :

$$\text{im}(\phi) = \{h \in G' \mid \exists g \in G \text{ such that } \phi(g) = h\}$$

It is easy to verify that $\text{im}(\phi)$ is a subgroup of G' .

It is furthermore neat to notice that the kernel of any homomorphism is a normal subgroup. We will shortly show the converse – that every normal subgroup is the kernel of some homomorphism.

Theorem 0.1. Let G and G' be groups and $\phi : G \rightarrow G'$ a group homomorphism. Then $\ker(\phi)$ is a normal subgroup of G .

Proof. Let $g \in G$ and $h \in \ker(\phi)$. We need to show that the kernel is closed under conjugation or equivalently: $ghg^{-1} \in \ker(\phi)$. This follows easily from definitions:

$$\phi(ghg^{-1}) = \phi(g)\phi(h)\phi(g^{-1}) = \phi(g)e\phi(g^{-1}) = e$$

□

The so-called "quotient" construction comes up frequently in the study of groups. Let us describe it now.

Definition 0.7. Let G be a group, H a subgroup and a an element of G . The set $aH = \{ah \mid h \in H\}$ is called a *left coset*. The set $Ha = \{ha \mid h \in H\}$ is called a *right coset*.

It is easy to see that the left cosets of a subgroup partition the group. From this it follows that they are disjoint sets. We now come to the first quotient construction.

Definition 0.8. Let G be a group and H a normal subgroup. The set of all left cosets of H is a group called the *quotient of G with H* and is denoted by G/H . The group law for combining two cosets is naturally defined by: $(aH)(bH) = (ab)H$. All of the group properties are easy to verify.

An easy example is the quotient of the group of real numbers with the integers: $R/Z = \mathbb{S}^1$. We have required H to be a normal subgroup in the above definition. In fact, when H is not normal, G/H is no longer a group. Nevertheless, it is still an important set and we call it a *quotient space*. An easy example of a quotient space is the 2-Sphere since $SO(3)/SO(2) = \mathbb{S}^2$.

We describe an important homomorphism from a group to the quotient with one of its subgroups.

Definition 0.9. Let G be a group and H a subgroup. Let $\phi : G \rightarrow G/H$ map each element $a \in G$ to the coset aH . Then ϕ is a homomorphism called the canonical quotient map.

It is now easy to see the converse of theorem 0.1 – that every normal subgroup is the kernel of some homomorphism. Given a group G and a normal subgroup H , one can verify that H is the kernel of the canonical quotient map.

Lastly, we introduce a method to combine two groups into a new one via a type of product.

Definition 0.10. Let G and H be groups and $\theta : H \rightarrow \text{Aut}(G)$ a homomorphism. Let $G \rtimes H$ be the set $G \times H$ with the following binary operation:

$$(g, h)(g', h') = (g[\theta(h)g'], hh') \quad (19)$$

It can be shown that $G \rtimes H$ is a group with identity element (e, e') and $(g, h)^{-1} = (\theta(h^{-1})(g^{-1}), h^{-1})$. The group $G \rtimes H$ is called the *semidirect product* of G and H . As stated in [15].

Appendix B: Elementary Topology

Definition 0.1. Let X be a set. Let \mathcal{T} be a collection of subsets of X called open sets such that $\emptyset, X \in \mathcal{T}$; a finite intersection of open sets is open; and an arbitrary union of open sets is open. A set X together with the collection \mathcal{T} is called a *topological space* and written as a pair (X, \mathcal{T}) .

Abuses of notation are frequent and bad, but unfortunately so prevalent that they are used here as well. A topological space (X, \mathcal{T}) will be written as X . The particular topology on X will be made clear in context. Let us give some examples of topologies.

Example 0.1. Let $X = \mathbb{R}$ be the set of real numbers. We define a subset S of X to be open if for each point $x \in S$ there exists some open interval T that contains x and is a subset of S . With this definition of open sets, the real line forms a topological space. This topology is called the *ordinary topology*.

Example 0.2. Let G be a topological group (see definition 2.2). Let H be a subgroup of G and ϕ the canonical quotient map from G to G/H . We call a set $S \subseteq G/H$ open if $\phi^{-1}(S)$ is open in G . The topology thus induced on G/H is called the *quotient topology*.

Example 0.3. Metric spaces with the standard definitions of open sets are topological spaces.

When dealing with metric spaces, we frequently come across some further terminology which we present here.

Definition 0.2. A *neighbourhood* of a point x is any set that contains an open set containing x .

Definition 0.3. Let X be a metric space. A sequence $\{x_n\}$ is called a *Cauchy sequence* if for any $\epsilon > 0$, there exists N such that for all $m, n \geq N$ we have that $d(x_n, x_m) < \epsilon$. (As stated in [16]).

Definition 0.4. A metric space is *complete* if every Cauchy sequence in it converges.

Definition 0.5. Let X be a metric space and $U \subseteq X$. We say that U is a *dense* subset of X if U along with the limits of all the sequences in U is equal to X . Alternatively, U is *dense* in X if the closure of U is equal to X .

Returning to the discussion of topological spaces, we define continuous functions. It is clear from the definition that continuous functions depend on the underlying topology. Therefore, a single set can have many different continuous functions depending on which topology it is endowed with.

Definition 0.6. Let X and Y be topological spaces and $f : X \rightarrow Y$ a map. If for all open sets $V \subseteq Y$, the set $f^{-1}(V)$ is open in X then we say that f is a continuous function.

Finally, we review the definition of compactness.

Definition 0.7. Let X be a topological space. Let $F = \{S_i\}_{i \in I}$ be a family of subsets of X . If every $S_i \in F$ is an open set, and the union of all the members in F is equal to X , then we call F an *open cover*. Furthermore, any subset of F , the union of whose members is equal to X , is called a *open subcover* of X .

Definition 0.8. Let X be a topological space. We say that X is *compact* if every open cover contains a finite subcover.

Appendix C: Global Spherical Signal Processing

The purpose of this appendix will be to present a number of continuous techniques available for signal processing on the sphere. Although it will be impossible to survey all of them, we present two frequently used ones – the Spherical Fourier Transform and spherical convolution. The existence of two tools on the sphere that are among the most important ones commonly available on Euclidean domains speaks to a rising interest in signal processing on the sphere. This demonstrates a need for generalization of wavelet-like transforms to the sphere and provides sufficient motivation (in the opinion of the author) for the derivation of the spherical continuous wavelet transform presented in Chapter 7.

1 Spherical Fourier Transform

As the name suggests, the Spherical Fourier Transform (SFT for short) is the spherical analogue of the traditional Fourier Transform in Euclidean space. With advent of fast algorithms in the last ten years through work of [19] and others, computing the SFT has become feasible. As a result, researchers have used the SFT in a wide variety of applications ranging from compression of human head models [7] and nonrigid shape recovery [17] to surface representations [5].

It is fairly easy to describe the SFT, for it is a projection of functions in $\mathcal{L}^2(\mathbb{S}^2)$ onto the set of spherical harmonics. One can invert the transform as a consequence of theorem 3.1 which states that spherical harmonics are an orthonormal basis for $\mathcal{L}^2(\mathbb{S}^2)$. Let us write this down formally. For any $f(\eta) \in \mathcal{L}^2(\mathbb{S}^2)$, the SFT is:

$$SFT_{lm} = \int_{\eta \in \mathbb{S}^2} f(\eta) \bar{Y}_{lm}(\eta) d\eta.$$

By the theorem above, we have a natural inversion formula:

$$f(\eta) = \sum_{l \in \mathbb{Z}^+, |m| \leq l} SFT_{lm}(\eta) Y_{lm}(\eta)$$

Let us now present an application of the spherical Fourier transform. This discussion will both demonstrate the usefulness of the SFT and at the same time emphasize that the SFT is a global and not a local technique in



Figure 2: From left to right: Reconstruction using 75 spherical harmonic coefficients. Reconstruction using 150 spherical harmonic coefficients.

the sense described above. The application in focus will be lossy 3D surface compression using spherical harmonics as described in [5].

The method is applicable to star-shaped surfaces and works as follows. Three-dimensional data is gathered by a device such as a range scanner and a unit sphere is fitted to the data collected. Next, the data is represented as a spherical function $f(\phi, \theta)$ whose value for each data point is its distance from the center of the sphere. This is the function that we wish to compress. The idea will be to compress $f(\phi, \theta)$ by projecting it onto a finite subset of the basis of spherical harmonics. To do this we apply the SFT to $f(\phi, \theta)$ and retain the first n largest coefficients. This is the "compressed" form of our data. At any point an approximation of the original data can be restored through the Inverse Spherical Fourier Transform (ISFT for short).

In terms of applicability, the main question is how efficient is this method? In other words, how large does n have to be (i.e how many coefficients does one need) to obtain a good approximation of the original data after reconstruction? Naturally, we would like to use a small number of coefficients and still get a fairly decent approximation. Consider the following two images.

The model was scanned in using a range laser scanner and the the original image had 75,000 points. The first image is an approximation of the model made using the above method and 75 spherical harmonic coefficients for reconstruction. We can see, that although global features are approximated well, local details are hard to make out and are sometimes erroneous. It turns out that by increasing the number of coefficients, the reconstruction is not much better. The reconstruction in the second image uses 150 spherical harmonic coefficients and still suffers from the same problem. One possible solution to this problem would be to use an different integral transform different from the SFT, one that is well-suited for the local analysis of signals on the sphere.

2 Convolution

In this section we discuss another important technique available in spherical signal processing — convolution. On the real line, convolution is an operation that quantifies the amount of overlap generated when a function $f(x)$ is shifted over another function $g(x)$. The convolution of these two functions is denoted by $f * g$ and given by the following integral:

$$f * g = \int_{-\infty}^{+\infty} f(\tau)g(x - \tau)d\tau \quad (20)$$

Convolution has a large number of application in signal processing including the important convolution theorem stating that the convolution of two functions is equal to the product of the Fourier transforms of these functions. Another use of convolution is to apply filters to signals defined on various domains. In the case of the plane, convolution can be used to apply a filter to any image defined on the plane. Similar applications arise on other domains such as the sphere and the definition of convolution can be extended to such domains in a number of ways.

The reason that in other domains there is a choice of how to define convolution is that there is a choice in how to define the "shifting" of one function over another. On the line shifting has only one possible interpretation — translation on the x -axis. However, on the plane we could define shifts by only by translations in the x and y directions or by translations in these directions as well as rotations about some axis. For non-isotropic functions this will make a difference.

Let us now discuss convolution on the sphere. For a proper treatment, let us first give the most general definition of a convolution over some locally compact group G equipped with a left Haar measure dx .

Definition 2.1. Let G be a locally compact group and $f, g \in \mathcal{L}^2(G, dx)$ be two complex valued measurable functions. Then the convolution of f and g denoted by $f * g$ is defined almost everywhere by:

$$(f * g)(x) = \int_{y \in G} f(y)g(y^{-1}x)dy \quad (21)$$

Notice that if $G = \mathbb{R}$, the above definition reduces to equation (20) — the usual definition of convolution on the real line. Moreover, notice that if G is not a commutative group, then $f * g \neq g * f$.

Recall that for a locally compact group that the map $x \rightarrow xy$ is continuous. This allows us to rewrite equation (21) in another equivalent way:

$$\begin{aligned}
(f * g)(x) &= \int_{y \in G} f(y)g(y^{-1}x)dy = \int_{y \in G} f(xy)g((xy)^{-1}x)dy \\
&= \int_{y \in G} f(xy)g(y^{-1}x^{-1}x)dy = \int_{y \in G} f(xy)g(y^{-1})dy
\end{aligned}$$

Therefore, there are two ways to define spherical convolution. In the first, we define "shifting" of a function $f(\eta)$ to be rotations around the θ and ϕ axes as well as rotation about the function itself. This corresponds to $SO(3)$ acting on $f(\eta)$. The corresponding definition is the following:

Definition 2.2. Let $f, g \in \mathcal{L}^2(\mathbb{S}^2, d\mu)$. Recall the definition of the rotation operator Λ in equation (1) Then the convolution of two functions is defined by:

$$(f * g)(\eta) = \int_{y \in SO(3)} f(y)\Lambda(y)g(\eta)d\mu \quad (22)$$

Using this definition, Driscoll and Healy [6] have proven an analogue of the convolution theorem for the sphere. For any square-integrable function f on the 2-sphere, let \hat{f} denote the SFT of f . Then the spherical convolution theorem says:

Theorem 2.1. Let $f, g \in \mathcal{L}^2(\mathbb{S}^2, d\mu)$. Then the following equality holds:

$$SFT(f * g)_{lm} = 2\pi \sqrt{\frac{4\pi}{2l+1}} \hat{f}_{l0} \hat{g}_{l0}$$

An alternative definition of convolution assumes that shifts are translations only along the θ and ϕ directions. It makes sense to use it when dealing with isotropic signals.

References

- [1] J-P. Antoine. Coherent states, wavelets and beyond. *World Scientific*, 2000.
- [2] J-P. Antoine and P. Vandergheynst. Wavelets on the 2-sphere: A group-theoretical approach. *Applied Computational and Harmonic Analysis*, 7:262–291, 1999.
- [3] Bernard W. Banks. Representations of the poincaré group.
- [4] Paul Bourdon, Wade Ramey, and Sheldon J. Axler. *Harmonic Function Theory*. Springer Verlag, 2001.
- [5] T. Bülow and K. Daniilidis. Surface representations using spherical harmonics and gabor wavelets on the sphere. *Technical Report MS-CIS-01-37*, 2001.
- [6] J.R. Driscoll and D.M. Healy. Computing fourier transforms and convolutions on the 2-sphere. *Advances in Applied Mathematics*, 15:202–250, 1994.
- [7] Sarp Ertük and Tim J. Dennis. Efficient representations of 3d human head models. *BMVC*, pages 329–339, 1999.
- [8] Gerald B. Folland. *A Course in Abstract Harmonic Analysis*. Springer-Verlag, 1995.
- [9] William Fulton and Joe Harris. *Representation Theory: A First Course*. Springer Verlag, 1991.
- [10] A. Grossman, J. Morlet, and T. Paul. Transforms associated to square integrable group representations. i: General results. *Journal of Mathematical Physics*, 26:2473–2479, October 1985.
- [11] A. Grossman, J. Morlet, and T. Paul. Transforms associated to square integrable group representations. ii: Examples. *Annales de L’Institut Henri Poincaré*, 45(3):293–309, 1986.
- [12] Paul R. Halmos. *Measure Theory*. Springer-Verlag, 1974.
- [13] Edwin Hewitt and Kenneth A. Ross. *Abstract Harmonic Analysis*. Springer Verlag, 1963.

- [14] M. Holschneider. *Wavelets: An Analysis Tool*. Oxford University Press, 1995.
- [15] Thomas W. Hungerford. *Algebra*. Springer-Verlag, New York, 1974.
- [16] Serge Lang. *Real and Functional Analysis*. Springer Verlag, 1993.
- [17] A. Matheny and D.B. Goldgof. The use of three and four-dimensional surface harmonics for rigid and nonrigid shape recovery and representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:959–971, October 1996.
- [18] Artin Michael. *Algebra*. Prentice Hall, 1991.
- [19] Martin J. Mohlenkamp. A fast transform for spherical harmonics. *Journal of Fourier Analysis and Applications*, 5:159–184, 1999.
- [20] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill, Inc., 1987.
- [21] G. Sansone. *Orthogonal Functions*. Dover Publications Inc., 1991.
- [22] Barry Simon. *Representations of Finite and Compact Groups*. American Mathematical Society, 1991.
- [23] G. Warner. *Harmonic Analysis on Semi-simple Lie Groups. I and II*. Springer, Berlin, 1972.